REMARKS

Claims 1-24, 27, 29-32, 34-46 and 212-216 were pending prior to this response, with claims 27, 39-32, 34-40 and 212-216 being withdrawn due to a restriction requirement. By the present communication, claims 27, 29-32, 34-40 and 212-216 have been canceled without prejudice, and claims 1-3, 11-14, 17, 19-21, 41, 43, and 44 have been amended to define Applicants' invention with greater particularity. Applicants respectfully request entry of the amendments set forth in this response under 37 CFR §1.116. The amendments do not raise any issues of new matter and the amended claims do not present new issues requiring further consideration or search. Support for the phrase "genomic DNA" in claim 1 is found, among others, at page 20, lines 3-10 and support for the phrase "in a liquid phase" may be found at page 23, lines 11-17. Support for the phrase "probe that comprises a DNA sequence that directs the synthesis of a biomolecule" may be found, among others, at page 53, lines 1-10. Support for the phrase "directs the synthesis" may further be found at page 16, lines 18-27, at page 44, lines 26-31, and at page 46, lines 19-25. Accordingly, claims 1-24 and 41-46 are currently pending.

The Rejection under 35 U.S.C. § 112, First Paragraph

Applicants respectfully traverse the rejection of claims 1-24 and 41-46 under 35 U.S.C. § 112, first paragraph, as failing to comply with the written description requirement due to allegedly lacking basis for the amendment to the claims to recite "naturally occurring polynucleotides." In particular, the Examiner alleges that the Specification fails to teach the concept of detecting "only naturally occurring polynucleotides." (Office Action, page 5) and requests citation of support for the amendment. Applicants submit that the Specification describes the collection and use of "naturally occurring polynucleotides" as follows:

> Culture-independent approaches to directly clone genes encoding both target enzymes and other bioactive molecules from environmental samples are based on the construction of libraries which represent the collective genomes of naturally occurring organisms, archived in cloning vectors that can be propagated in E. coli,

> Streptomyces, or other suitable hosts . Because the cloned DNA is initially extracted directly from environmental samples containing a mixed population of organisms, the representation of the libraries is not limited to the small fraction of prokaryotes that can be grown in pure culture, nor is it biased towards a few rapidly growing species.

(Specification, page 20, lines 3-10). Thus, Applicants respectfully submit that use of the term "naturally occurring polynucleotides" does not constitute addition of new matter to the Specification. However, to reduce the issues and advance prosecution, claims 1, 3, 11, 12, 14, 41, 43 and 44 have been amended to delete the phrase "naturally occurring polypeptides" and substitute the phrase "genomic DNA" to indicate that the polynucleotides obtained from the organisms are unmodified pieces of genomic DNA.

Accordingly, Applicants submit that the rejection of the claims for alleged new matter is now moot, and reconsideration and withdrawal of the rejection is respectfully requested.

The Rejection under 35 U.S.C. § 112, Second Paragraph

Applicants respectfully traverse the rejection of claims 1-24 and 41-46 under 35 U.S.C. § 112, second paragraph for allegedly being indefinite.

A. With regard to claim 43, the Examiner has maintained the rejection for indefiniteness over the phrase "encodes a small molecule" as used therein. In support of the rejection, the Examiner asserts: "The term small is a relative term, yet the claim does not set forth what the molecule is small in comparison to" (Office Action, page 5). Applicants maintain that the term "small" is not used to refer to a specific size of a molecule. The phrase "small molecule" is an art-recognized term to distinguish a chemical molecule or complex, such as a non-proteinaceous enzyme, from molecules containing amino acids or nucleic acids, either of which may be smaller in terms of molecular weight than a large chemical complex. Attached as Exhibit A are references supporting the notion that the phrase "small molecule" is an art-recognized term, and that one skilled in the art would understand the metes and bounds of the claimed subject matter.

B. With regard to claims 1-24 and 41-46, the Examiner asserts that the recitation of "naturally occurring polynucleotide" as used in claim 1 is indefinite because the polynucleotide in a library must be removed from a cell source, ligated to a vector and then inserted into a host cell, and the like. To clarify the meaning and intent of claim 1, Applicants have deleted the phrase "naturally occurring polynucleotide" and use instead the phrase "genomic polypeptides", which indicates wild type polynucleotides or unmodified genomic fragments. "Genomic DNA" may optionally be cut into fragments and inserted into vectors as described in the Specification. Support for this amendment is found in the Specification at page 20, lines 3-10.

In view of the above amendments, Applicants submit that the claims now meet all requirements under 35 U.S.C. § 112, second paragraph, and reconsideration and withdrawal of the rejection are respectfully requested.

The Rejection Under 35 U.S.C. § 102(e)

Applicants respectfully traverse the rejection of claims 1-5, 15, 16, 19-27, and 41-46 as allegedly being anticipated under 35 U.S.C. § 102(e) by Thompson et al. (U.S. Patent No. 5,824,485; hereinafter "Thompson"). Applicants submit that the invention methods for identifying a genomic polynucleotide that encodes a biomolecule having an activity of interest in a liquid phase, as defined by amended claim 1, distinguish over the disclosure of Thompson by requiring:

a) contacting in a liquid phase genomic DNA derived from one or more organisms with at least one nucleic acid probe that comprises a DNA sequence that directs the synthesis of a bioactivity or biomolecule having an activity of interest under conditions that allow hybridization of the probe to the genomic DNA; and

b) identifying genomic DNA that directs the synthesis of a biomolecule having an activity of interest in a host cell with an analyzer that detects DNA to which a probe has hybridized.

Applicants submit that the Examiner's argument in the present Office Action to the effect that Thompson uses hybridization probes to screen polynucleotides ligated into an expression vector (Office Action, page 10) ignores the fact that Thompson fails to teach that such a step can be used (i.e., without additional steps) to identify genomic DNA that directs the synthesis of a biomolecule, such as a small molecule or a protein having an activity of interest. For example, the Examiner refers to Thompson making and using expression libraries that contain a cDNA or genomic DNA fragment ligated to an *expression construct* (at Thompson col. 46 and 59 and claim 19) (Office Action, page 10). However, in the invention methods, there is no need for an expression library because the genomic DNA is detected without its expression (i.e., by its complementarity to the probe used in the invention method). Thompson is silent regarding detection of genomic DNA that directs the synthesis of a biomolecule having an activity of interest using a hybridization probe.

Accordingly, Applicants respectfully submit that since Thompson fails to disclose each and every element of amended claim 1 (and claims dependent thereon), *prima facie* anticipation under 35 U.S.C. 102(e) is not established over Thompson. Reconsideration and withdrawal of the rejection are, therefore, respectfully requested.

The Rejection under 35 U.S.C. § 103(a)

A. Applicants respectfully traverse the rejection of claims 1-11, 14-16, 19-24 and 41-46 under 35 U.S.C. 103(a) as allegedly being unpatentable over Thompson in view of Blumenfeld (U.S. Patent No. 6,228,580; hereinafter "Blumenfeld"). Applicants' remarks above regarding the deficiencies of Thompson for disclosing the invention methods apply equally and are incorporated here. In addition, Applicants submit that Thompson fails to suggest the invention methods because Thompson's focus in using hybridizing probes is to reduce the number of clones in a library by eliminating DNA that encodes genes related to primary organism functions in a gene subtraction method. In fact, Thompson is absolutely silent regarding hybridization

screening in a liquid phase of genomic DNA obtained from a mixed sample or for using a hybridizing probe (e.g., a labeled probe) to identify genomic DNA that directs the synthesis of a protein or small molecule having an activity of interest. Use of probes to remove genomic DNA that pertains to irrelevant primary organism functions leads away from using a probe to identify genomic DNA that directs the synthesis of a target biomolecule.

The Examiner relies upon Blumenfeld as disclosing the use of nucleic acid hybridization probes with a length of 100 to 1000 nucleotides or larger and asserts that it would have been obvious to use Blumenfeld's probes in the method of Thompson "to increase the specificity of hybridization and therefore increase the specificity of the method of detecting a polynucleotide having an activity of interest". However, there is no suggestion in the combined disclosures of Blumenfeld and Thompson to label a hybridization probe of greater than 100 nt for the purpose of identifying genomic DNA that directs the synthesis of a protein or small molecule having an activity of interest. Applicant urges the Examiner to consider the invention "as a whole" as required by the statute, rather than focusing on a single step of the invention as if it were not a step in an overall strategy.

Therefore, Applicants submit that the combined disclosures of Thompson and Blumenfeld do not suggest and would not motivate those of skill in the art to arrive at Applicants' methods as defined by amended claim 1, to identify in a liquid phase genomic DNA that directs the synthesis of a biomolecule having an activity of interest, such as protein or small molecule. Therefore, Applicants respectfully submit that *prima facie* obviousness of any pending claims is not established over the combined disclosures of Thompson and Blumenfeld. Consequently, reconsideration and withdrawal of the rejection over Thompson in view of Blumenfeld under 35 U.S.C. § 103(a) are respectfully requested.

B. Applicants respectfully traverse the rejection of claims 1-10, 13, 14, 16, 18-24 and 41-46 under 35 U.S.C. 103(a) as allegedly being unpatentable over Thompson in view of Hefti (U.S. Patent No. 6,340,568; hereinafter "Hefti"). Applicants' remarks above regarding the deficiencies

of Thompson for disclosing or suggesting the invention methods apply equally and are incorporated here.

The Examiner relies upon Hefti to cure the deficiencies of Thompson described above, particularly with respect to use of nucleic acid probes "up to 10,000 nucleotides in length in the method of Thompson" and detection of "molecular binding events" without the need for labeling the probe using multipole coupling spectroscopy. However, Hefti does not overcome the deficiencies of Thompson described above for suggesting to or for motivating those of skill in art to arrive at the invention methods because Hefti is absolutely silent regarding screening a library of genomic DNA in a liquid phase using hybridizing probes to identify genomic DNA *that encodes a protein or small molecule* having an activity of interest. Although Hefti is interested in MCS as a means of avoiding steric hindrance caused by the presence of the label in detection of dielectric properties, Hefti's focus is in determining the degree of hybridization or near hybridization using the extremely technical methods of MCS. Thus, Applicants respectfully submit that the combined disclosures of Thompson and Hefti do not establish *prima facie* obviousness of the invention methods under 35 U.S.C. § 103. Accordingly, reconsideration and withdrawal of the rejection over Thompson in view of Hefti under 35 U.S.C. § 103(a) are respectfully requested.

C. Applicants respectfully traverse the rejection of claims 1-17, 19-28, 34 and 41-46 under 35 U.S.C. 103(a) as allegedly being unpatentable over Thompson in view of Blumenfeld and Baselt (U.S. Patent No. 5,981,297). Applicants' remarks above regarding the deficiencies of the combined disclosures of Thompson and Blumenfeld for disclosing or suggesting the invention methods under 35 U.S.C. § 103(a) apply equally and are incorporated here.

The Examiner relies upon Baselt to cure the deficiencies of the combined disclosures of Thompson and Blumenfeld described above, particularly with respect to use of "labeling nucleic acid probes with magnetic molecules and detecting hybridization of probes to polynucleotides

having an activity of interest using SQUID" (Office Action, page 16), for example to increase speed and sensitivity of the assay.

However, Applicants submit that Baselt does not overcome the deficiencies of Thompson and Blumenfeld described above for suggesting to, or motivating those of skill in art to arrive at, the invention methods, as defined by amended claim 1, for screening genomic DNA in a liquid phase to identify genomic DNA that directs the synthesis of a biomolecule, such as a protein or chemical compound that has a desired bioactivity. Baselt is silent regarding identification of genomic DNA *that encode a protein or small molecule having an activity of interest.* Baselt's focus is to use a magnetic field detector, such as SQUID, to determine the presence of multiple analytes, or the concentration of an analyte in a liquid or gas phase and discloses binding molecules (e.g., probes) bound to the sensors. .

Therefore, Applicants submit that the combined disclosures of Thompson, Blumenfeld and Baselt do not suggest, and would not motivate, those of skill in the art to arrive at Applicants' methods of using hybridizing nucleic acid probes to screen a library of genomic DNA in a liquid phase to identify those that direct the synthesis of a protein or small molecule having a bioactivity having an activity of interest. Accordingly, Applicants submit that *prima facie* obviousness of any pending claims is not established under 35 U.S.C. § 103(a) over the combined disclosures of Thompson, Blumenfeld, and Baselt, and reconsideration and withdrawal of the rejection are respectfully requested.

Conclusion

In view of the above amendments and remarks, reconsideration and favorable action on claims 1-24 and 41-46 is respectfully requested. In the event any matters remain to be resolved, the Examiner is requested to contact the undersigned at the telephone number given below so that a prompt disposition of this application can be achieved.

Enclosed is Check No. 573576 in the amount of $760.00; which consists of $250.00 for the appeal fee and $510.00 for Three (3) Month Extension of Time fee. The Commissioner is hereby authorized to charge any other fees that may be associated with this communication, or credit any overpayment, to Deposit Account No. 07-1896.

Respectfully submitted,

Date: January 21, 2005

Lisa A. Haile, J.D., Ph.D.
Registration No.: 38,347
Telephone: (858) 677-1456
Facsimile: (858) 677-1465

DLA PIPER RUDNICK GRAY CARY US LLP
4365 Executive Drive, Suite 1100
San Diego, California 92121-2133
USPTO CUSTOMER NO. 28213

EXHIBIT A

RNA as a target for small molecules

Steven J Sucheck and Chi-Huey Wong*

Proteins are folded to form a small binding site for catalysis or ligand recognition and this small binding site is traditionally the target for drug discovery. An alternative target for potential drug candidates is the translational process, which requires a precise reading of the entire mRNA sequence and, therefore, can be interrupted with small molecules that bind to mRNA sequence-specifically. RNA thus presents itself as a new upstream target for drug discovery because of the critical role it plays in the life of pathogens and in the progression of diseases. In this post-genomic era, RNA is becoming increasingly amenable to small-molecule therapy as greater structural and functional information accumulates with regard to important RNA functional domains. The study of aminoglycoside antibiotics and their binding to 16S ribosomal RNA has been a paradigm for our understanding of the ways in which small molecules can be developed to affect the function of RNA.

Addresses
Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA
*e-mail: wong@scripps.edu

Current Opinion in Chemical Biology 2000, 4:678–686

1367-5931/00/$ – see front matter

Abbreviations

CFTR	cystic fibrosis transmembrane conductance regulator
HIV-1	type 1 human immunodeficiency virus
HTS	high-throughput screening
RNP	ribonucleoprotein
RRE	Rev response element
TAR	*trans*-activating region

Introduction

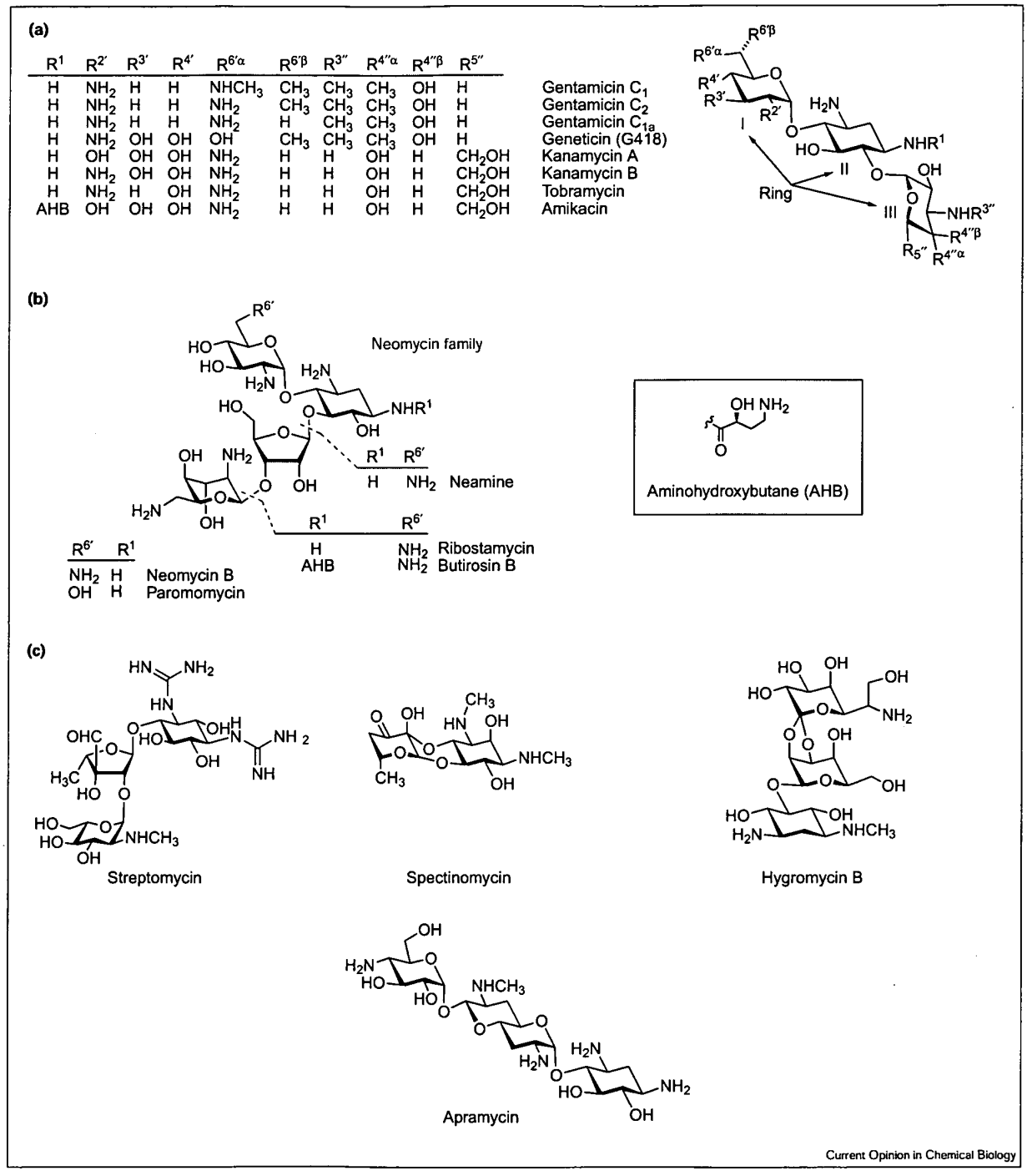
The potential of RNA as a therapeutic target for small-molecule drug discovery continues to improve as the relationships between RNA structure and function become increasingly apparent in the lifecycles of organisms and in the development of disease processes. This potential for small-molecule drug discovery has not gone unnoticed. Recently, the topic has been addressed in multiple reviews [1–3]. The primary function of RNA in the life of the cell is in protein synthesis. In this process, RNA serves as a template (mRNA), ribosome component (rRNA) and an activated intermediate (i.e. aminoacyl-tRNA). RNA also participates in the expression of genes by catalyzing the maturation of mRNAs via ribozymes. Considering these essential cellular roles, it is not surprising that the functions of RNA are highly regulated by numerous ribonucleoprotein (RNP)–RNA interactions [4]. Thus, protein synthesis, mRNA maturation and RNP–RNA interactions are generally considered the most promising points of intervention for the discovery and development of small molecule effectors of RNA. Within the past decade, the number of RNA-based molecular targets has begun to grow rapidly as the detailed structural and functional relationships of RNA, particularly rRNA, have been elucidated [5–8]. Our understanding of the nature by which small molecules can affect the function and structure of RNA continues to improve in parallel with advances in the technology for generating and obtaining structural information on small-molecule–RNA complexes [9–12]. Analysis of these structures has resulted in rapid evolution in the design of new therapeutically useful RNA-binding ligands.

Potential of RNA as a drug target

Although RNA has only recently been viewed as a target for small-molecule drug discovery, the advantages of targeting RNA over traditional protein targets are quickly being realized. The potential for the slower development of drug resistance against small molecules is one example. RNA functional domains are more highly conserved and perhaps more accessible than the shapes of enzyme active sites. Thus, it is expected that pathogens will find it difficult to mutate their RNA and develop resistance. Type 1 human immunodeficiency virus (HIV-1), which has rapidly developed resistance to enzyme inhibitors, illustrates this point. RNA-based antiviral targets offer a potential solution to the problem. The *trans*-activating region (TAR) RNA, responsible for gene regulation in HIV-1, has been identified as a possible RNA-based target [13,14]. The RNA functional domain of TAR binds the cognate peptide Tat, which activates transcription of the HIV-1 genome. Aminoglycosides (Figure 1), a class of structurally diverse aminocyclitols with potent antibiotic activity were the first small molecules shown to disrupt the TAR–Tat interaction [15,16]. Since these initial findings, other classes of small molecules have been found to effect more potent and selective inhibition of this system. Recently, a Tat agonist based on acridine, CGP40336A, achieved an IC_{50} of 22 nM for inhibition of Tat binding (Figure 2; [17,18]). The 2,4-diaminoquinozalines and quinoxaline-2,3-diones have also been shown to selectively and stoichiometrically bind TAR in the nanomolar range (Figure 2; [16]). In model systems, these small molecules show inhibition of the growth of the virus [16]. Therefore, these small molecules serve as the prototypical inhibitors of protein–RNA and peptide–RNA interactions.

Unique biochemical effects such as enhanced gene expression might also be achieved by the disruption of RNP–RNA or RNA–RNA interaction [4]. The cystic fibrosis transmembrane conductance regulator (CFTR) illustrates another example of the role a small molecule can play in effecting the expression of a gene. Point mutations in CFTR that introduce a stop codon have been shown to cause severe cystic fibrosis in a 5% subpopulation of cystic fibrosis patients. It is

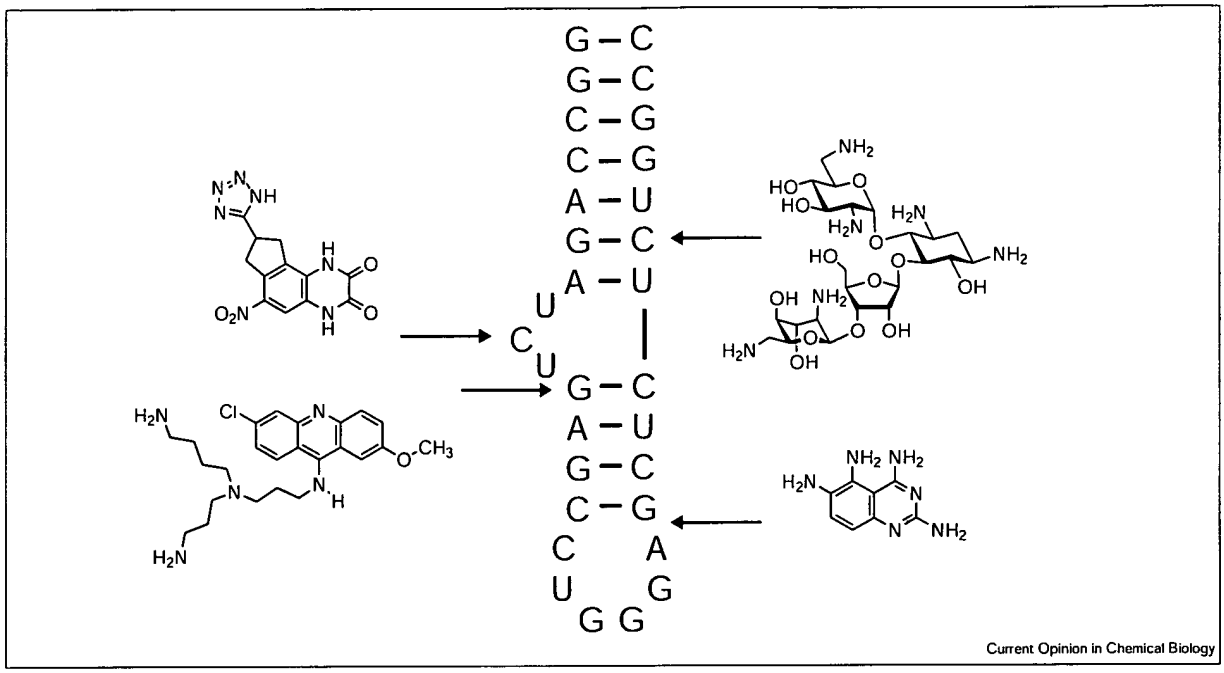
Figure 1



Several representative structures of the streptamine and deoxystreptamine aminoglycoside antibiotics. **(a)** Gentamycin and kanamycin aminoglycosides. **(b)** The neomycin family. **(c)** Other aminoglycosides.

well known that aminoglycosides such as G-418 and gentamicin bind 16S rRNA and interfere with proof reading during protein synthesis [19–21]. The addition of these aminoglycosides to an *in vitro* protein-expression system for

Figure 2

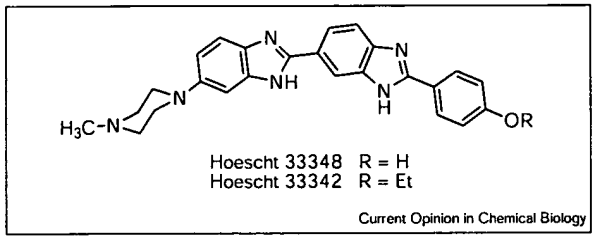


These small molecules, neomycin B (top right; [15]), the 2,4-diaminoquinozalines (bottom right) and the quinoxaline-2,3-diones (top left; [17,18]), as well as the acridine-based compound CGP40336A (bottom left; [16]), bind HIV-1 TAR RNA and antagonize the binding of Tat. The arrows indicate the binding sites on TAR.

mutant CFTR promoted read-through of stop codon [22]. This aminoglycoside-induced read-through resulted in expression levels of more than 35% of wild-type CFTR. The results suggest that 16S rRNA represents an important target for the treatment of certain patients.

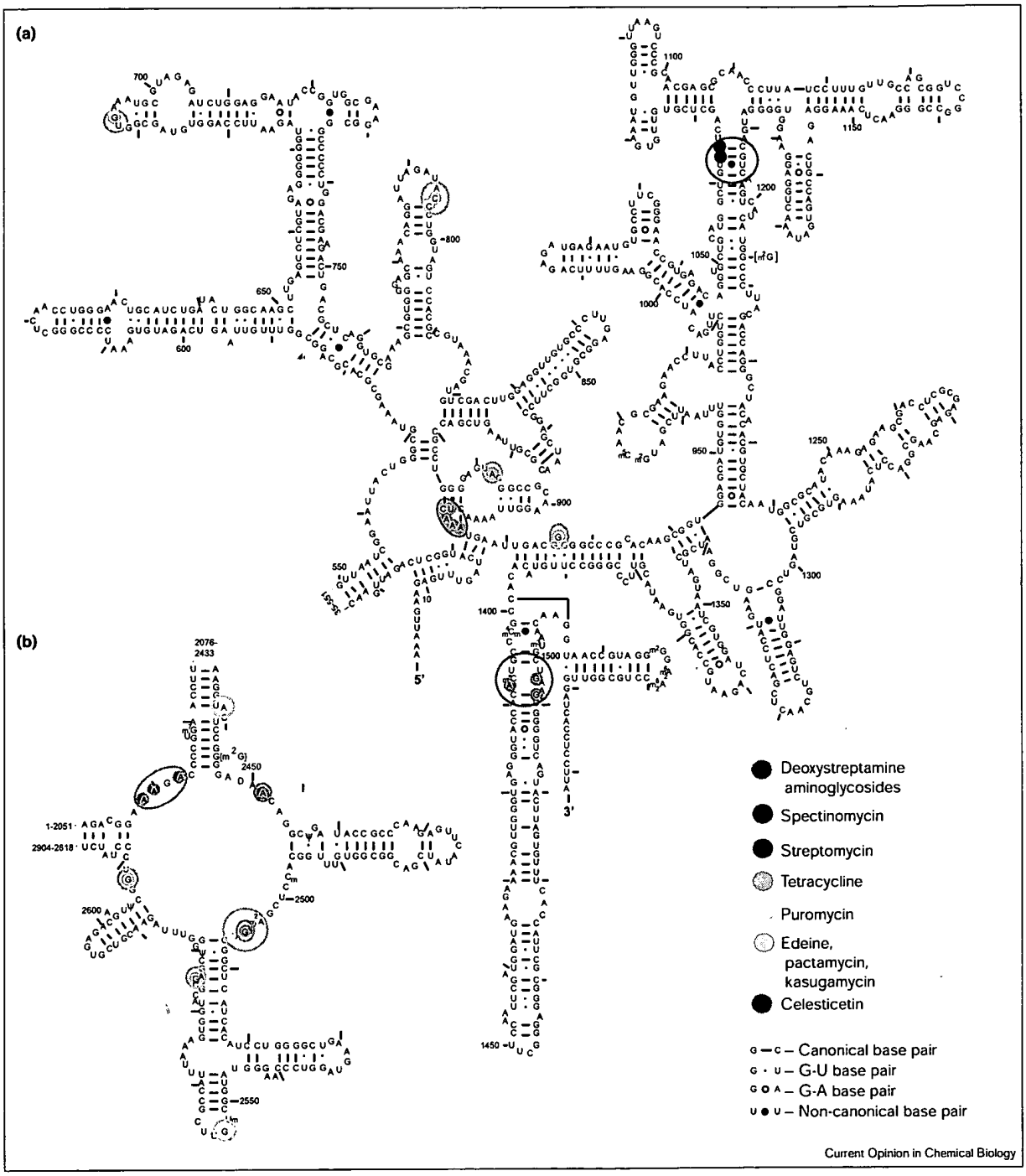
In a landmark report by Werstuck and Green [23], engineered genes containing kanamycin A and tobramicin aminoglycoside-binding aptamers were transfected into bacteria to give rise to aminoglycoside-resistant phenotypes. The significance of these experiments was to demonstrate that small-molecule RNA-binding motifs were capable of discriminating between closely related molecules. The results are impressive because aminoglycosides are highly charged and are promiscuous in their interaction with RNA at higher concentrations. The authors further demonstrated that small molecules could be used to regulate gene expression. RNA aptamers for Hoescht 33348 (Figure 3) were generated and inserted into the 5′-untranslated region of a mammalian β-galactosidase (β-gal) expression plasmid. Addition of Hoescht 33342, closely related to Hoescht 33348 (Figure 3), reduced β-gal expression by 90% without affecting luciferase, the control gene [23]. These studies indicate that small molecules may prove to be valuable tools for the study of genes, particularly with respect to their role in cellular processes. It remains to be seen whether or not small-molecule regulation of gene expression can be used to target a particular gene at will without prior knowledge of a specific RNP–RNA interaction or detailed information regarding a functional domain in the RNA sequence. The use of a small aminoglycoside library to target the PAX3-FKHR and Bcr-Abl [24,25] oncogenes, which are responsible for certain forms of alveolar rhabdomyosarcoma and leukemia, respectively, demonstrated this strategy. Both of these genes arise from chromosomal translocations that have unique breakpoint sequences not found in the wild-type genes.

Figure 3



The structures of Hoescht 33348 and 33342. Hoescht was shown to selectively inhibit the expression of a β-galactosidase gene when a high affinity RNA aptamer for Hoescht 33348 was inserted in the 5′-untranslated region of the gene.

Figure 4

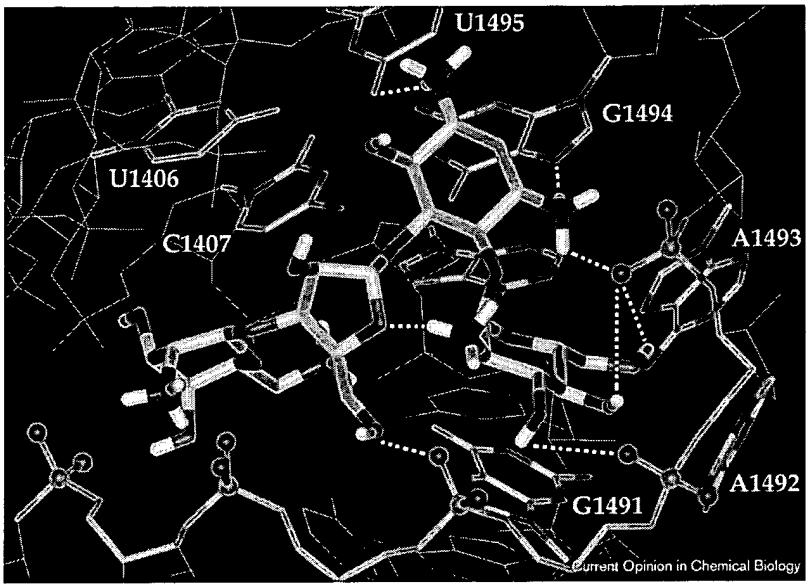


The secondary structures of *E. coli* **(a)** 16S and **(b)** 32S rRNA. The binding sites for several antibiotics are indicated.

The results of these studies suggest that it is feasible to find small molecules that bind selectively to breakpoints in oncogenes [26].

Other potential RNA-based targets include RNA viruses such as influenza virus, hepatitus C and HIV. For example, a current potential target for antiviral therapy is the mRNA

Figure 5



Solution structure of paromomycin complexed with the 16S RNA model sequence. Dotted lines indicate putative hydrogen bonds. Hydrogen atoms on carbon have been omitted for clarity. See also [41].

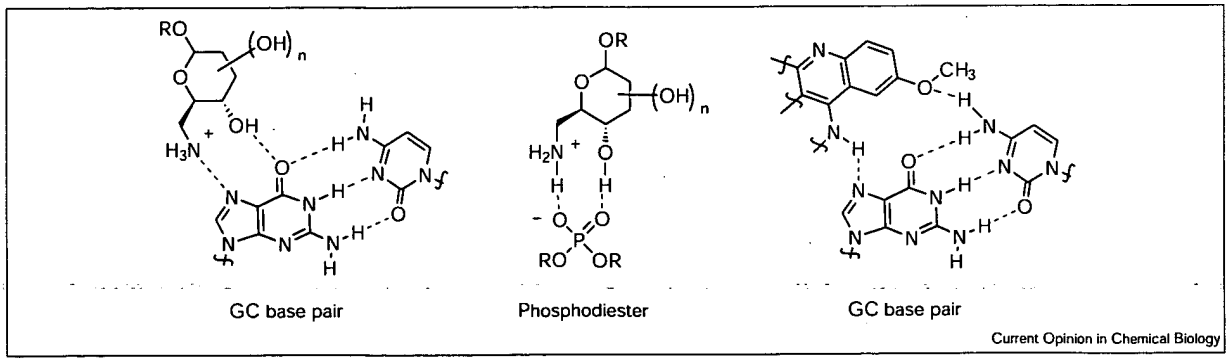
encoding influenza hemagglutinin protein (HA), which is essential for viral infection and is responsible for the pathogenicity of influenza A [27]. Other proteins with HA-like activation mechanisms are the HIV-1 Env protein [28] and the Ebola GP_1 protein [29]. In addition, it is generally difficult to develop specific inhibitors to target the enzymes or receptors that share a common substrate (e.g. ATP for kinases) or ligands; however, targeting the corresponding RNA sequence-specifically may be a feasible alternative, as illustrated in the inhibition of oncogenic RNA [26]. Thus, targeting the RNA of these proteins may afford an alternative mechanism for treatment of numerous diseases.

Finally, technological advances in high-throughput screening (HTS) hold the promise of simplifying the identification of selective RNA-binding small molecules. Recently, a prototype system, using the aminoglycosides 16S A-site rRNA system as a model, showed that small-molecule–RNA binding can be detected by electrospray mass spectroscopy, which can be adapted to HTS [30,31]. The recent development of gene chip technology is also likely to have a profound impact on small-molecule RNA screening. As chip technology matures, chips with the capacity to hold entire genomes are expected [32]. This technology is a prime candidate for adaptation to the selection of small-molecule RNA effectors.

Small-molecule–RNA interactions

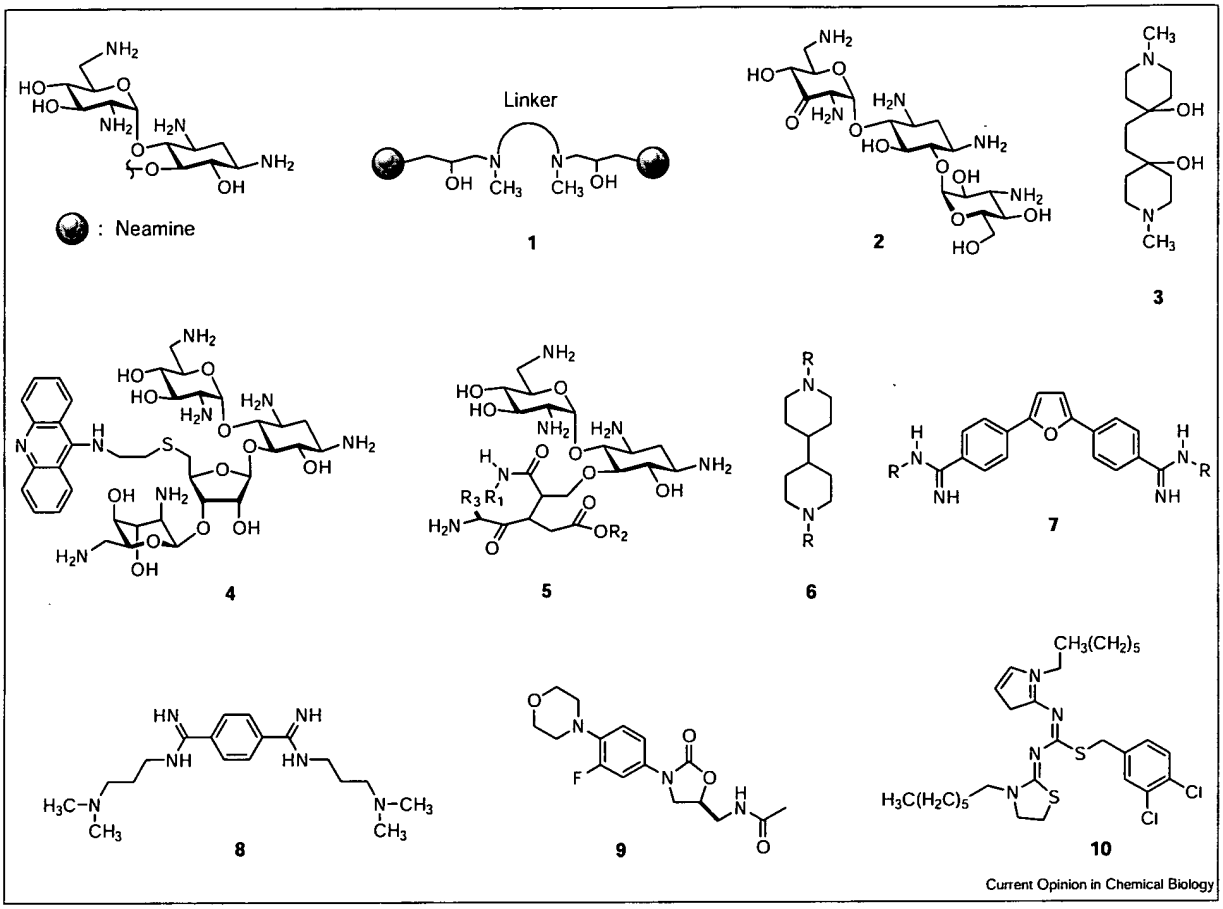
The study of small molecule RNA effectors has primarily focused on the aminoglycosides (Figure 1). These

Figure 6



Putative modes of interaction for hydroxyamines with phosphodiesters, and the Hoogsteen faces of GC base pairs.

Figure 7

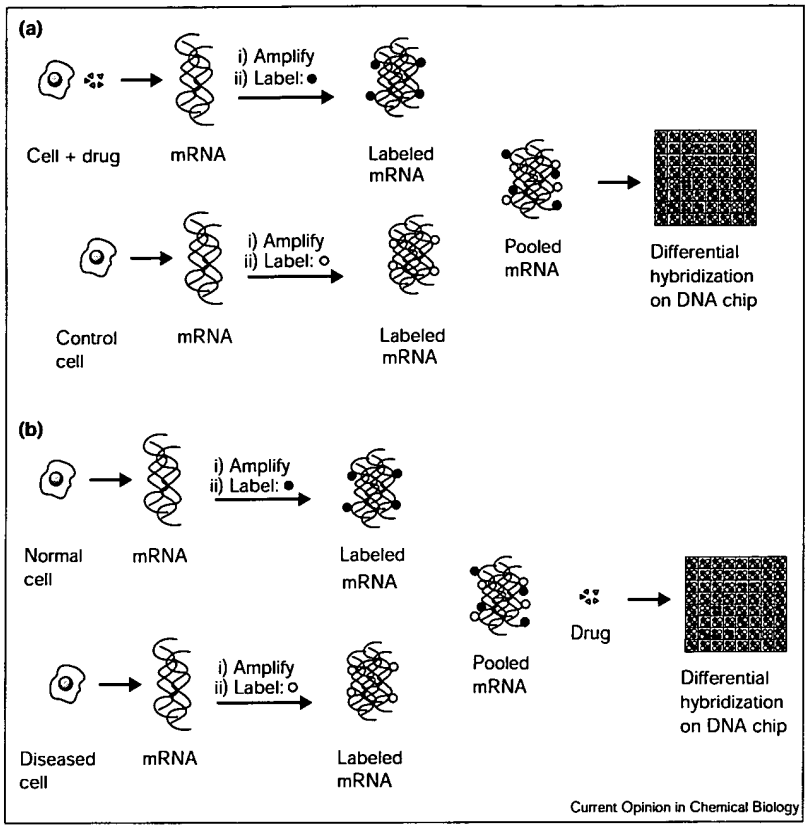


Various aminoglycoside and non-aminoglycoside-based effectors of RNA function. For 16S rRNA: neamine dimer **1** [50], enzymaticly modified kanamycin A analogue **2** [61], aminol **3** [62]; HIV-1 RRE RNA: acridine-neomycin B conjugate **4** [63], *O*-5-modified neamine library **5** [48], 4,4′-bipiperidine derivatives **6** (R = various guanidinium groups) [58], diphenyl furans **7** [64]; HIV-1 frameshifting RNA: DB213 **8** [65]; 50S RNA: Linzolid **9** [66]; GTPase rRNA: thiostrepton **10** [67].

compounds were among the fist to be recognized as effectors of RNA function. Aminoglycoside–RNA interactions have been particularly well-defined for 16S [33,34] and 23S [35–37] ribosomal RNA (Figure 4). The antibiotic activity of these compounds is believed to result from the effect these molecules have on the translational accuracy of protein synthesis [19–21]. The nanomolar binding affinity and surprising selectivity that aminoglycosides have for their target RNA has made the study of these compounds with RNA the paradigm for small-molecule–RNA interactions [38–40]. Detailed structural studies of the aminoglycosides paromomycin and gentamicin C_{1a}, with RNA sequences, modeling the A-site domain of 16S RNA (Figure 5; [41,42]), have provided significant insight into the molecular motifs required to achieve selective RNA recognition. The biophysical studies of this system indicate that aminoglycoside-binding to 16S RNA induces a conformational change that stabilizes the structure of 16S rRNA [43]. The effect of A-site stabilization on protein synthesis is slower dissociation of the tRNA–16S-rRNA complex [19], which prevents efficient proof reading during protein synthesis. These findings have shaped our view on small-molecule effectors of RNA.

On the basis of the study of 16S rRNA with aminoglycosides and other small molecules, we conclude that the structure of RNA is dynamic and this is fundamental to its functional features. The conformational changes in RNA induced by aminoglycoside binding the 16S A-site appear to represent a general mechanism for the activity for many small-molecule RNA effectors. The aminoglycoside antibiotic spectinomycin, which induces a conformational change in helix 34 of 16S RNA, is another example of this mechanism of action [44]. In this case, the effect of conformational change on protein synthesis is inhibition of A to P-site translocation, which occurs by preventing elongation factor G from binding the active conformation of helix 34 [45]. Similar conformational

Figure 8



Gene chip technology. **(a)** Traditional DNA gene chip approach for finding new enzyme targets. **(b)** Potential application of gene chip technology for the screening of small molecules that effect the expression of undesirable genes.

change effects have been observed by circular dichroism studies using neomycin B binding to the TAR element of HIV-1 RNA [46]. The resulting conformational change in TAR increases the rate of dissociation of the cognate peptide, Tat, which results in the antiviral activity of this compound. Aminoglycosides have also been used to establish a precedent for inhibition of the HIV-1 Rev–Rev response element (RRE) interaction [47,48].

The study of aminoglycoside–16S-RNA complexes has improved our understanding of the determining factors in RNA recognition. The 16S rRNA A-site model shows the 2′-amino-2′-deoxyglucose moiety, ring I, of the aminglycoside in a stacking interaction with an adjacent base, thus stabilizing the entire structure of the hairpin. The modeling studies also suggest that the 1,3-hydroxylamine motif, present in most aminoglycosides, is an important RNA recognition motif for chelating phosphate and the edges of nucleotide bases in RNA. Furthermore, evidence from the study of 6′-amino-6′-deoxyglucose supports this thesis [49]. The other amines and hydroxyl moieties in aminoglycosides form hydrogen bonds and salt bridges with phosphate and the edges of RNA bases as well. Important RNA binding motifs are illustrated in Figure 6. The generation of RNA aptamers, which have been used to find new aminoglycoside–RNA binding motifs, has also been a useful tool for the understanding of these interactions [9,11]. The stabilization of RNA structure in highly selective small-molecule–RNA complexes is generally observed when a small rigid molecule binds a region of RNA that has formed a binding pocket, which are created by bulges in the RNA. Recently, new aminoglycoside-based ligands for 16S RNA that bind by a bivalent interaction have emerged. The design of these structures resulted from an understanding of the important RNA recognition motifs present in the aminoglycosides as well as careful study of the secondary binding sites present in 16S RNA [50]. Related strategies have been employed for the development of high-affinity ligands for L-21 *Sca* I ribozyme form *Tetrahymena* [51].

Further study of aminoglycosides and their interaction with RNA has revealed that they are capable of interacting with a number of other functional RNA domains. Some of these include human immunodeficiency virus [15,46], group I introns [52], the hammerhead ribozyme [53] and hepatitis delta virus [54]. The study of these systems has resulted in the development of new classes of small molecules with high selectivity for target

sequences, increased efficacy and improvement in their pharmacokinetic character [55–58]. Selected structures are illustrated in Figure 7.

Advances in high-throughput screening for small molecule–RNA interactions

The study of the interaction between aminoglycosides and RNA has formed a basis for the development of new HTS technologies for the discovery of small-molecule–RNA interactions. Screens based on fluorescence and surface plasmon resonance permit screening of small-molecule RNA-binding libraries [59,60]. Furthermore, new NMR and mass spectrometry methods have the capacity to screen tens of thousands of small-molecule–RNA complexes per day [30,31]. The recent advancements in HTS on oligonucleotide microarrays may also be of great benefit for identifying novel small-molecule–RNA interactions. With the capacity to hold several million genes, this technology is currently being used to drive the identification of new therapeutic targets. The current approach for finding new enzyme-based targets is to grow cells in the presence of a drug candidate. The genes affected by the drug are identified and the protein products of these genes are investigated for their potential as new drug targets (Figure 8a). This technology has the potential to be applied to the discovery of novel small molecule RNA binders as well. For example, small molecules might be used to interfere with mRNA–DNA hybridization as a means of identifying novel small-molecule–RNA interactions (Figure 8b). mRNA from normal and diseased organisms could be differentially labeled. Looking for changes that occur only in the expression of diseased mRNA while not affecting mRNA expression in normal cells, could be used to identify highly selective effectors of RNA function. These promising new technologies are expected to accelerate the discovery of new classes of RNA-binding ligands.

Conclusions

Recent studies on RNA structures and RNA–small-molecule interactions have provided insights into the molecular basis of RNA recognition. It appears that, from the drug development standpoint, targeting RNA may have some advantages as compared with targeting proteins. For example, more sites are accessible at the RNA level, whereas the 'active site' of a protein is often the only target. Proteins that share a common substrate (e.g. ATP for kinases) or ligand are, in general, difficult to inhibit specifically, but targeting the RNA that encodes the protein of interest sequence-specifically has been demonstrated. In addition, development of multivalent drugs to target RNA, or drugs that target the RNA sequence essential for encoding an important sequence of a protein in order to tackle the problem of drug resistance or affect the proteins function is also feasible. With a greater understanding of the genomic information from different species, targeting RNA with small molecules is becoming a new frontier in drug discovery.

Acknowledgement

We thank the National Institutes of Health for support of this research.

References

1. Hermann T, Westhof E: **RNA as a drug target: chemical, modelling, and evolutionary tools.** *Curr Opin Biotechnol* 1998, **19**:66-73.
2. Tor Y: **RNA and the small molecule world.** *Angew Chem Int Ed* 1999, **38**:1579-1582.
3. Ecker DJ, Griffey RH: **RNA as small-molecule drug target: doubling the value of genomics.** *Drug Discov Today* 1999, **4**:420-429.
4. Hermann T: **Strategies for the design of drugs targeting RNA and RNA-protein complexes.** *Angew Chem Int Ed* 2000, **39**:1890-1905.
5. Yuan RC, Steitz JA, Moore PB, Crothers DM: **The 3′ terminus of 16S rRNA: secondary structure and interaction with ribosomal protein S1.** *Nucleic Acids Res* 1979, **7**:2399-2418.
6. Ban N, Nissen P, Hansen J, Capel M, Moore PB, Steitz TA: **Placement of protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit.** *Nature* 1999, **400**:841-847.
7. Batey RT, Rambo RP, Doudna JA: **Tertiary motifs in RNA structure and folding.** *Angew Chem Int Ed* 1999, **38**:2327-2343.
8. Agalarov Sultan C, Prasad GS, Funke PM, Stout CD, Williamson JR: **Structure of the S15, S6, S18-rRNA complex: assembly of the 30S ribosome central domain.** *Science* 2000, **288**:107-112.
9. Wang Y, Rando RR: **Specific binding of aminoglycoside antibiotics to RNA.** *Chem Biol* 1995, **2**:281-290.
10. Harada K, Martin SS, Frankel AD: **Selection of RNA-binding peptides *in vivo*.** *Nature* 1996, **380**:175-179.
11. Jiang L, Suri AK, Fiala R, Patel DJ: **Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex.** *Chem Biol* 1997, **4**:35-50.
12. Ramos A, Gubser CC, Varani G: **Recent solution structures of RNA and its complexes with drugs, peptides and proteins.** *Curr Opin Struct Biol* 1997, **7**:317-323.
13. Gait MJ, Karn J: **RNA recognition by the human immunodeficiency virus Tat and Rev proteins.** *Trends Biochem Sci* 1993, **18**:255-259.
14. Churcher MJ, Lamont C, Hamy F, Dingwall C, Green SM, Lowe AD, Butler JG, Gait MJ, Karn J: **High affinity binding of TAR RNA by the human immunodeficiency virus type-1 Tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region.** *J Mol Biol* 1993, **230**:90-110.
15. Mei H-Y, Galan AA, Halim NS, Mack DP, Moreland DW, Sanders KB, Truong HN, Czarnik AW: **Inhibition of an HIV-1 Tat-derived peptide binding to TAR RNA by aminoglycoside antibiotics.** *Bioorg Med Chem Lett* 1995, **5**:2755-2760.
16. Mei H-Y, Cui M, Heldsinger A, Lemrow SM, Loo JA, Sannes-Lowery KA, Sharmeen L, Czarnik AW: **Inhibitors of protein-RNA complexation that target the RNA: specific recognition of human immunodeficiency virus type 1 TAR RNA by small organic molecules.** *Biochemistry* 1998, **37**:14204-14212.
17. Hamy F, Brondani V, Floersheimer A, Stark W, Blommers MJJ, Klimkait T: **A new class of HIV-1 Tat antagonist acting through Tat–TAR Inhibition.** *Biochemistry* 1998, **37**:5086-5095.
18. Gelus N, Hamy F, Bailly C: **Molecular basis of HIV-1 TAR RNA specific recognition by an acridine Tat-antagonist.** *Bioorg Med Chem* 1999, **7**:1075-1079.
19. Cabanas MJ, Vazquez D, Modolell J: **Dual interference of hygromycin B with ribosomal translocation and with aminoacyl-tRNA recognition.** *Eur J Biochem* 1978, **87**:21-27.
20. Misumi M, Nishimura T, Komai T, Tanaka N: **Interaction of kanamycin and related antibiotics with the large subunit of ribosomes and the inhibition of translocation.** *Biochem Biophys Res Commun* 1978, **84**:358-365.
21. Campuzano S, Vazquez D, Modolell J: **Functional interaction of neomycin B and related antibiotics with 30S and 50S ribosomal subunits.** *Biochem Biophys Res Commun* 1979, **87**:960-966.
22. Howard M, Frizzel RA, Bedwell DM: **Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations.** *Nature Med* 1996, **2**:467-469.

23. Werstuck G, Green MR: **Controlling gene expression in living cells through small molecule-RNA interations.** *Science* 1998, **282**:296-298.

24. Shtivelman E, Lifshitz B, Gale RP, Canaani E: **Fused transcript of Abl and Bcr genes in chronic myelogenous leukemia.** *Nature* 1985, **315**:550-554.

25. Raitano AB, Whang YE, Sawyers CL: **Signal transduction by wild-type and leukemogenic Abl proteins.** *Biochim Biophys Acta* 1997, **1333**:F201-F216.

26. Sucheck SJ, Greenberg WA, Tolbert TJ, Wong C-H: **Design of small molecules that recognize RNA: development of aminoglycosides as potential antitumor agents that target oncogenic RNA sequences.** *Angew Chem Int Ed* 2000, **39**:1080-1083.

27. Steinhauer DA: **Role of hemagglutinin cleavage for the pathogenicity of influenza virus.** *Virology* 1999, **258**:1-20.

28. McCune JM, Rabin LB, Feinberg MB, Lieberman M, Kosek JC, Reyes GR, Weissman IL: **Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus.** *Cell* 1988, **53**:55-67.

29. Volchkov VE: **Processing of the Ebola virus glycoprotein.** *Curr Top Microbiol Immunol* 1999, **235**:35-47.

30. Hofstadler SA, Sannes-Lowery KA, Crooke ST, Ecker DJ, Sasmor H, Manalili S, Griffey RH: **Multiplexed screening of neutral mass-tagged RNA targets against ligand libraries with electrospray ionization FTICR MS: a paradigm for high-throughput affinity screening.** *Anal Chem* 1999, **71**:3436-3440.

31. Griffey RH, Hofstadler SA, Sannes-Lowery KA, Ecker DJ, Crooke ST: **Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry.** *Proc Natl Acad Sci USA* 1999, **96**:10129-10133.

32. Lipshutz RJ, Fodor SPA, Gingeras TR, Lockhart DJ: **High density synthetic oligonucleotide arrays.** *Nat Genet* 1999, **21**:20-24.

33. Moazed D, Noller HF: **Interaction of antibiotics with functional sites in 16S ribosomal RNA.** *Nature* 1987, **327**:389-394.

34. Woodcock J, Moazed D, Cannon D, Davies J, Noller HF: **Interaction of antibiotics with A- and P-site-specific bases in 16S ribosomal RNA.** *EMBO J* 1991, **10**:3099-3103.

35. Egebjerg J, Garrett RA: **Binding sites of the antibiotics pactamycin and celesticetin on ribosomal RNAs.** *Biochimie* 1991, **73**:1145-1149.

36. Oehler R, Polacek N, Steiner G, Barta A: **Interaction of tetracycline with RNA: photoincorporation into ribosomal RNA of *Escherichia coli*.** *Nucleic Acids Res* 1997, **25**:1219-1224.

37. Rodriguez-Fonseca C, Phan H, Long KS, Porse BT, Kirillov SV, Amils R, Garrett RA: **Puromycin-rRNA interaction sites at the peptidyl transferase center.** *RNA* 2000, **6**:744-754.

38. Miyaguchi H, Narita H, Sakamoto K, Yokoyama S: **An antibiotic-binding motif of an RNA fragment derived from the A-site-related region of *Escherichia coli* 16S rRNA.** *Nucleic Acids Res* 1996, **24**:3700-3706.

39. Wang Y, Hamasaki K, Rando RR: **Specificity of aminoglycoside binding to RNA constructs derived from the 16S rRNA decoding region and the HIV-RRE activator region.** *Biochemistry* 1997, **36**:768-779.

40. Recht MI, Fourmy D, Blanchard SC, Dahlquist KD, Puglisi JD: **RNA sequence determinants for aminoglycoside binding to an A-site rRNA model oligonucleotide.** *J Mol Biol* 1996, **262**:421-436.

41. Fourmy D, Recht MI, Blanchard SC, Puglisi JD: **Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic.** *Science* 1996, **274**:1367-1371.

42. Yoshizawa S, Fourmy D, Puglisi JD: **Structural origins of gentamicin antibiotic action.** *EMBO J* 1998, **17**:6437-6448.

43. Fourmy D, Yoshizawa S, Puglisi JD: **Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA.** *J Mol Biol* 1998, **277**:333-345.

44. Brink MF, Brink G, Verbeet MP, de Boer HA: **Spectinomycin interacts specifically with the residues G1064 and C1192 in 16S rRNA, thereby potentially freezing this molecule into an inactive conformation.** *Nucleic Acids Res* 1994, **22**:325-331.

45. Bilgin N, Richter AA, Ehrenberg M, Dahlberg AE, Kurland CG: **Ribosomal RNA and protein mutants resistant to spectinomycin.** *EMBO J* 1990, **9**:735-739.

46. Wang S, Huber PW, Cui M, Czarnik AW, Mei H-Y: **Binding of neomycin to the TAR element of HIV-1 RNA induces dissociation of Tat protein by an allosteric mechanism.** *Biochemistry* 1998, **37**:5549-5557.

47. Zapp ML, Stern S, Green MR: **Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production.** *Cell* 1993, **74**:969-978.

48. Park WKC, Auer M, Jaksche H, Wong C-H: **Rapid combinatorial synthesis of aminoglycoside antibiotic mimetics: use of a polyethylene glycol-linked amine and a neamine-derived aldehyde in multiple component condensation as a strategy for the discovery of new inhibitors of the HIV RNA Rev responsive element.** *J Am Chem Soc* 1996, **118**:10150-10155.

49. Hendrix M, Alper PB, Priestley ES, Wong C-H; **Hydroxyamines as a new motif for the molecular recognition of phosphodiesters: implications for aminoglycoside-RNA interactions.** *Angew Chem Int Ed* 1997, **36**:95-98.

50. Sucheck SJ, Wong AL, Koeller KM, Boehr DD, Draker K-A, Sears P, Wright GD, Wong C-H: **Design of bifunctional antibiotics that target bacterial RNA and inhibit resistance-causing enzymes.** *J Am Chem Soc* 2000, **122**:5230-5231.

51. Michael K, Wang H, Tor Y: **Enhanced RNA binding of dimerized amino glycosides.** *Bioorg Med Chem* 1999, **7**:1361-1371.

52. von Ahsen U, Noller HF: **Footprinting the sites of interaction of antibiotics with catalytic group I intron RNA.** *Science* 1993, **260**:1500-1503.

53. Stage TK, Hertel KJ, Uhlenbeck OC: **Inhibition of the hammerhead ribozyme by neomycin.** *RNA* 1995, **1**:95-101.

54. Chia JS, Wu HL, Wang HW, Chen DS, Chen PJ: **Inhibition of hepatitis delta virus genomic ribozyme self-cleavage by aminoglycosides.** *J Biomed Sci* 1997, **4**:208-216.

55. Katja M, Tor Y: **Designing novel RNA binders.** *Chem Eur J* 1998, **4**:2091-2098.

56. Afshar M, Prescott CD, Varani G: **Structure-based and combinatorial search for new RNA-binding drugs.** *Curr Opin Biotechnol* 1999, **10**:59-63.

57. Chen Q, Shafer RH, Kuntz ID: **Structure-based discovery of ligands targeted to the RNA double helix.** *Biochemistry* 1997, **36**:11402-11407.

58. Li K, Fernandez-Saiz M, Rigl CT, Kumar A, Ragunathan KG, McConnaughie AW, Boykin DW, Schneider HJ, Wilson WD: **Design and analysis of molecular motifs for specific recognition of RNA.** *Bioorg Med Chem* 1997, **5**:1157-1172.

59. Hendrix M, Priestley ES, Joyce GF, Wong C-H: **Direct observation of aminoglycoside-RNA interactions by surface plasmon resonance.** *J Am Chem Soc* 1997, **119**:3641-3648.

60. Hamasaki K, Rando RR: **A high-throughput fluorescence screen to monitor the specific binding of antagonists to RNA targets.** *Anal Biochem* 1998, **261**:183-190.

61. Haddad J, Vakulenko S, Mobashery S: **An antibiotic cloaked by its own resistance enzyme.** *J Am Chem Soc* 1999, **121**:11922-11923.

62. Tok JB-H, Rando RR: **Simple aminols as aminoglycoside surrogates.** *J Am Chem Soc* 1998, **120**:8279-8280.

63. Kirk SR, Luedtke NW, Tor Y: **Neomycin-acridine conjugate: a potent inhibitor of Rev-RRE binding.** *J Am Chem Soc* 2000, **122**:980-981.

64. Ratmeyer L, Zapp M, Green M, Vinayak R, Kumar A, Boykin DW, Wilson WD: **Inhibition of HIV-1 Rev-RRE interaction by diphenylfuran derivatives.** *Biochemistry* 1996, **35**:13689-13696.

65. Hung M, Patel P, Davis S, Green SR: **Importance of ribosomal frameshifting for human immunodeficiency virus type 1 particle assembly and replication.** *J Virol* 1998, **72**:4819-4824.

66. Shinabarger D: **Mechanism of action of the oxazolidinone antibacterial agents.** *Expert Opin Invest Drugs* 1999, **8**:1195-1202.

67. Rogers MJ, Bukhman YV, McCutchan TF, Draper DE: **Interaction of thiostrepton with an RNA fragment derived from the plastid-encoded ribosomal RNA of the malaria parasite.** *RNA* 1997, **3**:815-820.

Chemistry & Biology, Vol. 9, 1275–1285, December, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S1074-5521(02)00284-3

Small Molecules, Big Impact: A History of Chemical Inhibitors and the Cytoskeleton

Review

Jeffrey R. Peterson[1] and Timothy J. Mitchison
Department of Cell Biology and
Institute of Chemistry and Cell Biology
Harvard Medical School
Boston, Massachusetts 02115

Chemical inhibitors, whether natural products or synthetic, have had an enormous impact on the study of the eukaryotic cytoskeleton. Here we review the history of some of the most widely used cytoskeletal poisons and their influence on our understanding of cytoskeletal functions. We then highlight several new inhibitors and the targeted screens used to identify them and discuss why this approach has been successful.

Introduction

Few problems in biology have seen such a strong impact from the development of small molecule tools as the study of cell morphogenesis and the subsequent elucidation of the underlying anatomy of the cell that we now call the cytoskeleton. Compounds targeting the cytoskeleton are among the most commonly used chemical inhibitors in basic cell biological research. In addition, several of these have been developed into bona fide drugs widely used in the treatment of such diseases as cancers and gout. The goal of this review is to briefly review the history of a few of the most common inhibitors of the cytoskeleton with an emphasis on the impact these molecules had and continue to have on this field. By "impact," we mean either a significant conceptual leap in our understanding or a novel technique that becomes widely used. We will then highlight a few recent examples of novel small molecule inhibitors identified in screens targeting the cytoskeleton and discuss the promise that chemical approaches offer for the future of research on the cytoskeleton.

Colchicine and the Identification of Tubulin

The identification of the target of colchicine as tubulin, the subunit comprising the ubiquitous microtubule cytoskeleton of cells, is a remarkable example of forward chemical genetics. Indeed, the discovery of tubulin is intimately tied to the identification of the colchicine target. This tropolone derivative, found in the meadow saffron (genus *Colchicum*), has been used medicinally since at least the 18th century (and continues to be used) in the treatment of gout, and it is widely used as a research tool for the study of microtubules. Only in 1940 was the structure of the active component, colchicine, determined [1], and by the 1950s the effects of colchicine had been studied in cells and tissues of many types (for a comprehensive review of the early history of colchicine, see [2]).

Early investigation of the cellular effects of colchicine described the "explosion" of mitotic figures observed in tissues of colchicine-treated plants and animals (reviewed in [3]). Although we now understand that this arises from the arrest of the normal cell division cycle in mitosis, it was initially considered that colchicine could be inducing an altered mitosis in treated tissues that was called "c-mitosis." In plants, colchicine proved a rapid and convenient tool to generate agriculturally important polyploid strains, quickly replacing previous methods such as heat shock or treatment with other chemicals (extensively reviewed in [2]). In addition, the increased prevalence of mitotic figures in colchicine-treated cells was used to unambiguously determine that 46 chromosomes is the normal human diploid number rather than the previously believed 48 [3, 4]. Thus, even prior to the identification of the mechanism of action of colchicine, it was widely used in the areas of medicine, agriculture, and biology.

The determination of tubulin as the protein target of colchicine by Ed Taylor and colleagues in the late 1960s stands as a landmark in the identification of small molecule targets in complex mixtures as well as opening up the microtubule field by identifying the protein subunit that comprises these filaments. Using radiolabeled colchicine prepared by methylation of colchiceine in tritiated water, Borisy and Taylor biochemically characterized a colchicine binding activity in both intact cells and cell extracts [5, 6, 7]. This binding activity was found to be enriched in cells and tissues containing abundant microtubules, suggesting that the target of colchicine was the subunit of microtubules. Taylor and colleagues subsequently purified the colchicine binding protein from both sperm tails and mammalian brain and characterized it as a 120 kDa dimer containing 2 moles of bound GTP, thus identifying the molecular subunit of microtubules [8, 9]. The name "tubulin" was provided by Mohri [10], who determined the amino acid composition of the sea urchin sperm microtubule subunit. Thus, colchicine was at the same time the agent for tying microtubules to important cellular processes such as mitosis and the agent of protein (gene) discovery, fulfilling the requirements of forward chemical genetics.

Taxol and Nocodazole

In 1971, a natural product with antileukemic and antitumor activity was identified from an alcohol extract of the bark of the western yew (*Taxus brevifolia*) and named taxol [11]. Progress on taxol lagged due to its perceived low antitumor activity, the limited quantities of the compound, and scarcity of the source tree [12, 13]. Nevertheless, later observation of cells isolated from taxol-treated mice revealed the presence of abnormal mitotic figures [14]. Remarkably, in contrast to other microtubule poisons (colchicine, nocodazole, the *Vinca* alkaloids, eg.), taxol was shown to *stimulate* the polymerization of microtubules both in vitro [15] and in vivo [16, 17]. With this discovery, then, two distinct natural products had been identified with opposing activities on microtubule stability.

[1]Correspondence: jeffrey_peterson@hms.harvard.edu

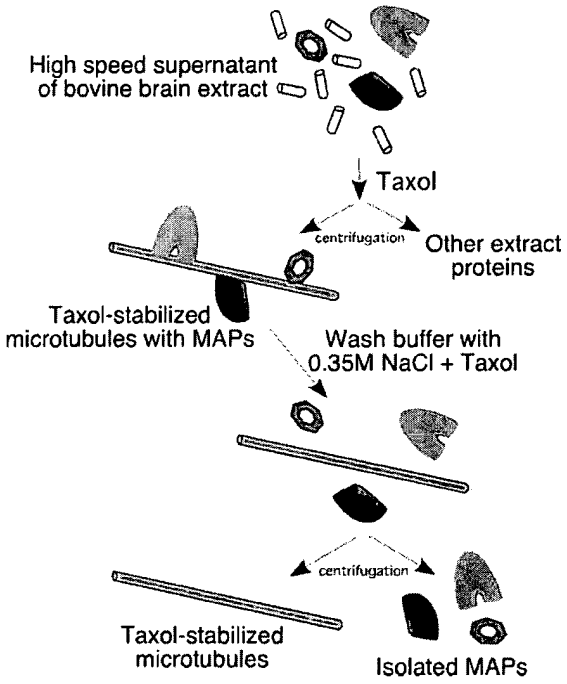


Figure 1. Use of Taxol-Stabilized Microtubules for the Isolation of MAPs

Richard Vallee isolated *m*icrotubule-*a*ssociated *p*roteins by adding taxol to induce the polymerization and stabilization of microtubules in a soluble extract of bovine brain. Taxol-stabilized microtubules were subsequently stripped of their associated proteins in a high-salt wash and pelleted, leaving the isolated MAPs in the supernatant [18].

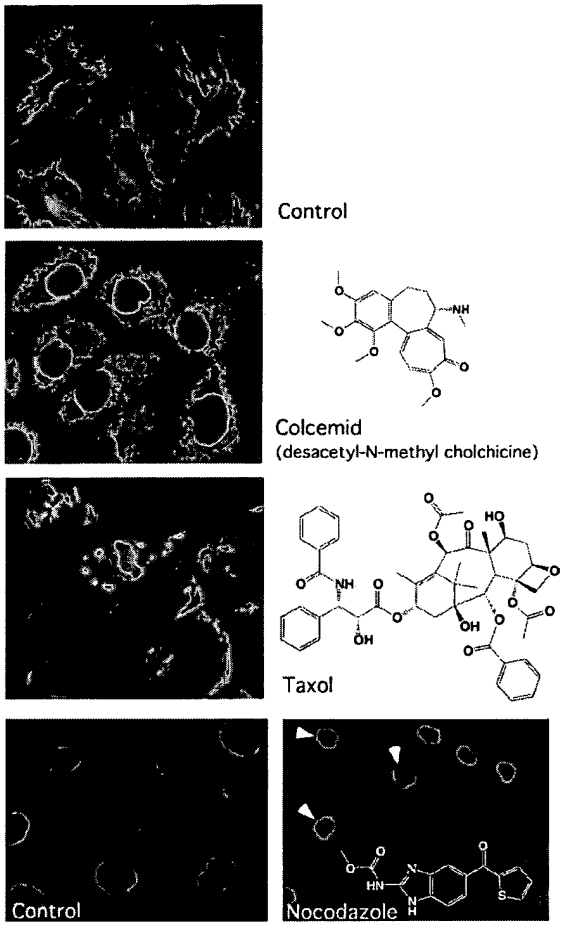


Figure 2. Structure and Effect of Classical Microtubule Poisons

In the upper panels, B-SC-1 cells treated with 2 μM colcemid or 10 μM taxol for 1 hr were fixed and stained with anti-tubulin antibodies (green), and DNA was stained with Hoechst (blue). In the lower panels, B-SC-1 cells treated for 24 hr with 3.3 μM nocodazole or vehicle control were stained with Hoechst. Note the increased prevalence of mitotic nuclei containing condensed DNA (several examples are marked with arrowheads).

The subsequent impact of taxol on basic biological research was dramatic. Vallee [18] exploited the strong stabilizing influence of taxol on microtubules to purify them and their bound microtubule-associated proteins (MAPs) from bovine brain (see Figure 1). Taxol was added to an extract of brain to polymerize microtubules and allow binding of endogenous MAPs. These filaments were then centrifuged and collected. A subsequent high-salt wash of the pellet stripped the MAPs, while the constant presence of taxol maintained the structural integrity of the microtubules, which could then be centrifuged away from the soluble MAPs. Thus, taxol allowed an affinity-based purification of MAPs that, because of the instability of microtubules to the high-salt extraction, would not have been possible otherwise. A similar microtubule affinity purification using taxol later aided the discovery and study of the microtubule-based motor protein kinesin [19]. Taxol-stabilized microtubules have also been used as the substrate to visualize gliding motility powered by both major microtubule-based motor families, kinesin and dynein, immobilized on glass coverslips.

A *synthetic* compound directly affecting microtubules was identified in Belgium (Janssen Pharmaceutica) in 1975 in a screen for antihelminthic compounds and was termed oncodazole (R 17934), presumably for an observed antitumor activity [20]. Two years later, this benzimidazole compound appears in the literature by the name nocodazole, and this term has endured to the present, perhaps because the early clinical promise against cancer was not realized. As with taxol, the microtubule disruption observed in nocodazole-treated cells led researchers to test directly if nocodazole bound to tubulin. De Brabander and colleagues showed that nocodazole indeed bound tubulin and in a manner competitive with colchicine [21].

What Have We Learned from Microtubule Poisons?

Immunofluorescence staining using anti-tubulin antibodies reveals a dramatic disruption of the microtubule network in cells treated with microtubule poisons (Figure 2). Nocodazole along with colchicine have been used to demonstrate functional roles for microtubules in numerous cell biological processes, including the anchoring of the Golgi complex at the microtubule-organizing center (reviewed in [22]), cell migration, and tumor inva-

sion (reviewed in [23]). Nocodazole is preferred over colchicine in basic research when reversible inhibition of microtubule polymerization is required. The dissociation rate of colchicine from tubulin is very slow [6], which helped in the identification of tubulin by Taylor and colleagues.

An additional role for microtubules uncovered with the help of microtubule-disrupting agents is the trafficking of intracellular particles. The microtubule cytoskeleton is generally arrayed from a central organizing center termed the MTOC from which microtubules radiate throughout the cell (see Figure 2). Melanophores utilize this microtubule "highway" for the intracellular distribution of thousands of pigment granules that migrate along this network using microtubule-based motors. The dispersion of granules to the cell periphery produces an apparent "darkening" in the skin, whereas the aggregation of granules at the MTOC produces the opposite effect. In 1965, prior to the identification of tubulin, Stephen Malawista [24] demonstrated that colchicine perturbed the aggregation of frog melanocyte granules, which he interpreted in terms of a "decreased protoplasmic viscosity." Once the target of colchicine had been identified, subsequent work using colchicine and the *Vinca* alkaloids helped demonstrate that microtubules play a fundamental role in the movement of pigment granules (reviewed in [25, 26]). The importance of microtubules to the intracellular transport of other organelles and vesicles remains an important question in cell biology.

The Cell Cycle, Checkpoints, and Cancer

Since taxol, nocodazole, and the colchicine relative colcemid all block normal chromosome segregation, Schimke and coworkers tested the affect of microtubule disruption by these agents on cell cycle progression of numerous mammalian cell lines [27]. The authors observed that human cell lines generally arrested in mitosis in compound-treated cells, consistent with the postulate of a checkpoint that ensures that cells remain in mitosis until a proper spindle is assembled [28]. Indeed, inhibition of cell cycle progression is one of the most prominent features of cells treated with microtubule poisons, although cell lines vary in the effectiveness of this arrest [29]. The mitotic arrest induced by these compounds also provides a convenient manner to synchronize the cell cycle state of cultured cells. Cultured cells treated with microtubule-depolymerizing agents accumulate in mitosis (see Figure 2, lower panels) and can be synchronously released by removing the compound from the media. This feature is now widely exploited by those studying mitosis for the isolation of cells homogeneously arrested in the mitotic state.

Benomyl, an agricultural fungicide and microtubule-polymerization inhibitor structurally related to nocodazole, was used in a similar manner to explore this checkpoint genetically in the yeast *Saccharomyces cerevisiae* [30, 31]. Both groups identified yeast mutants that did not properly arrest cell cycle progression in response to disruption of microtubules by benomyl. These mutants, termed "bub" for "budding uninhibited by benomyl" and "mad" for "mitotic arrest deficient," identified several genes critical for ensuring the proper temporal order of cell cycle events, and understanding their function remains a central question in cell biology.

Their usefulness in arresting cell division has led to the assessment of many microtubule inhibitors for the treatment of cancer. Indeed, the microtubule-destabilizing *Vinca* alkaloids vinblastine (originally vincaleukoblastine [32]) and vincristine helped establish a link between microtubules and cancer. These closely related but chemically distinct compounds from leaves of the Madagascar periwinkle were originally isolated based on their ability to depress white blood cell counts (reviewed in [33]). This original observation has now matured into the current use of vinblastine and vincristine in the clinical treatment of Hodgkin's lymphoma and leukemia, respectively. The development and history of microtubule poisons for clinical use is outside the scope of this review, however, and the interested reader is referred to any of a number of reviews (e.g., [32, 12, 34]). Nevertheless, it deserves mentioning that taxol, colchicine, and the *Vinca* alkaloids are mature, modern pharmaceuticals and, because of the central role of the cytoskeleton in cell division, cytoskeletal proteins remain important anticancer targets [35]. Indeed, the next generation of tubulin-targeting, anticancer compounds is being developed to address limitations of the current arsenal, such as aqueous solubility, multidrug resistance, and the pronounced toxicity toward lymphocytes and peripheral neurons. One well-characterized example, the natural product epothilone, stabilizes microtubules with greater potency than taxol, is less sensitive than taxol to P-glycoprotein-mediated multidrug resistance, and remains active against taxol-resistant tumor models (reviewed in [36, 37]). Several total syntheses of epothilone have been achieved, and medicinal chemistry efforts have identified derivatives with improved pharmacological properties (reviewed in [38, 39]). Readers interested in a more comprehensive and mechanistic view of small molecules that target tubulin and their anticancer potential are directed to other reviews (see [23, 40]). Those interested in recent developments in the chemistry of taxol are referred to [41].

Cytochalasins, Phalloidin, and the Actin Cytoskeleton

The family of mold metabolites known as cytochalasins were independently isolated from distinct fungal species by Aldridge et al. [42] at Imperial Chemical Industries Ltd. and by Rothweiler and Tamm at the University of Basel [43]. Whereas Rothweiler and Tamm called their compound Phomin after the *Phoma* species from which they isolated the compound, Carter [44] provided the name cytochalasin from the Greek *cytos* (a cell) and *chalasis* (relaxation) to describe the effects of this compound on mouse fibroblasts. A preliminary article based on the work at ICI appeared in New Scientist, calling this compound family "one of the most remarkable groups of biologically active substances to be described in years," although perplexingly the name was spelled "cytocholasins" [45]. Cytochalasin inhibited whole-cell migration, ruffling of the cell margin, and cytoplasmic cleavage of dividing cells, but nuclear division continued, thus producing multinucleated cells over time.

Carter, knowing nothing about the target of cytochalasin, used this compound to probe two important biological questions: how do cells migrate and how do cells divide? Having recently proposed a mechanism for cell motility based on surface tension and differential adhesion of cells to substrates, which he termed 'haptotaxis' [46], it is not surprising that he wrongly interpreted the effects of cytochalasin on cell motility in terms of the compound increasing the adhesivity of the cell membrane to the substrate, thus preventing both forward movement and cell ruffling. Nevertheless, using cytochalasin he was able to imply a common molecular mechanism underlying cell ruffling, motility, and cytokinesis. Furthermore, this motility was distinguishable from the movement of spermatozoa, ciliates, and flagellates, which were not affected by cytochalasin and are now known to be actin-independent phenomena. Cytochalasin represented the first compound that could disrupt cytokinesis (cell division) without affecting karyokinesis (nuclear division), thus clearly establishing the independence of karyokinesis from cytokinesis. The usefulness of a specific inhibitor for studying the mechanism of cytokinesis had already been anticipated by Wolpert [47].

In order to understand the impact of cytochalasin in this area, it is necessary to review contemporary theories on the mechanism of cell cleavage. Previously, numerous ideas had been proposed to explain cytokinesis (reviewed in [47]). One of these, the cortical gel contraction hypothesis [48], proposed that contraction of a cortical network underlying the deepening cleavage furrow between daughter cells provided the force for cytokinesis. This model stood in opposition to surface expansion theories which suggested that plasma membrane expansion was an active process providing the energy for furrow ingression [49] or theories in which the mitotic apparatus itself or other subcortical components are responsible for force generation [50, 51].

Despite its postulate of a cortical gel composed of interlinked "elongate protein components" that could undergo a "forcible folding without relinquishing their intermolecular linkages" during contraction [48], the cortical contraction theory lacked a morphological structure that could be pointed to as the source of force. Using electron microscopy, several investigators subsequently identified circumferential filaments underlying the cleavage furrow which where proposed to represent the apparatus of the 'contractile ring' ([52] and references therein). It was Schroeder, however, who, by demonstrating that cytochalasin both disrupted this contractile ring and abolished cytokinesis [52, 53], directly implicated the filaments in cytokinesis and suggested a "purse string" mechanism for furrowing. This work, of course, also suggested a mechanism of action for cytochalasin with respect to the block of cytokinesis.

Morphologically similar microfilaments had already been observed in nondividing cells [54], although their relationship to the filaments of the ring remained speculative [52]. Ishikawa et al. [55] used a technique for decorating muscle actin filaments for electron microscopy using a proteolytic fragment of the myosin protein to probe the nature of microfilaments in nonmuscle cells. The characteristic arrowhead pattern they observed in animal fibroblasts as well as epidermal and epithelial cells suggested that these nonmuscle microfilaments could be related to muscle actin. Similar observations in *Acanthamoeba* [56], epithelial brush border [57], *Dictyostelium* [58], *Physarum* [59], and *Amoeba* [60] supported the universiality of these structures. These studies, as well as the biochemical characterization of actin-like filaments derived from nonmuscle cells, supported the ubiquity of actin. Still, the connection between these microfilaments, the filaments of the cytokinetic furrow, and the actin filaments of muscle was tenuous.

Wessells and coworkers studied the effects of cytochalasin and the microtubule inhibitor colchicine, whose target had been recently identified, on axonal outgrowth of cultured neurons [61]. They noted that cytochalasin rapidly disrupted microspikes and growth cone dynamics, whereas colchicine only affected the axon and on a much slower time scale. These important observations suggested a strong connection between the microfilaments and growth cone motility while establishing an important but distinct role for the microtubule system in axonal outgrowth. Similarly, work using cytochalasin on the glands of the oviduct and salivary epithelium showed that disruption of the microfilament network strongly perturbed gland morphogenesis [62, 63]. Mounting evidence seemed to suggest that the microfilaments themselves were the target of cytochalasin:

> The following processes are sensitive to cytochalasin: cytokinesis, cell movement, axonal growth cone and microspike activity, maintenance and change in shape of salivary gland epithelium, formation and maintenance of early glands in oviduct, and perhaps migration of nuclei in an epithelium preparatory to mitosis. Every such case can be explained if contractile filaments are rendered inoperative by the drug; and in every case so far examined, morphological alterations in microfilaments have resulted from application of cytochalasin. ...The common sensitivity to cytochalasin suggests a homology between those filaments comparable to that between microtubules from varying cell types in their sensitivity to colchicine [63].

The final demonstration that cytochalasin targets the microfilaments directly arrived in 1972. Parallels between the contractile ring and contraction in muscle had been made much earlier (reviewed in [47]), yet no molecular connection had been made between the two beyond the observation that the actin 'thin filaments' of muscle were similar in diameter to the 'microfilaments' observed in cells [64].

In a landmark paper, Spudich and Lin studied the effect of cytochalasin on purified actin and actomyosin from rabbit muscle [64]. These authors demonstrated that cytochalasin could decrease the viscosity of solutions of pure filamentous actin, thus revealing in one fell swoop that cytochalasin targets the actin protein of muscle, and actin therefore likely also comprises the contractile ring and other cellular microfilaments. This supported the prevailing hypothesis that actomyosin assemblies controlled aspects of the motility of nonmuscle cells and led to an explosion in the use of cytocha-

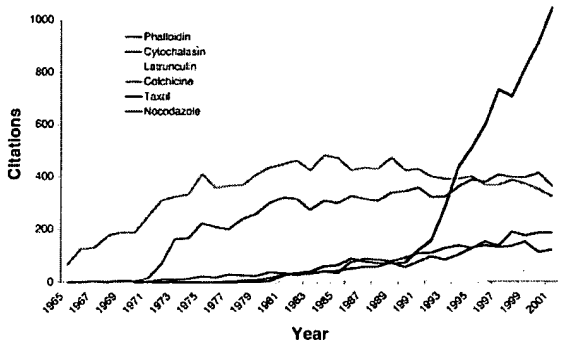


Figure 3. PubMed Citations of Cytoskeletal Inhibitors
The number of citations for each search term is plotted as a function of year.

lasin as a research tool for disrupting actin function (see Figure 3).

These studies connecting actin, microfilaments, and cytochalasin were reinforced by contemporary work on the fungal metabolite phalloidin. It was isolated in 1937 as one of the toxic peptides derived from mushrooms of the *Amanita* genus [65], whose toxicity has been studied since the early 1800s (reviewed in [66, 67, 68]). Morphological studies of intoxicated animals revealed only acute liver toxicity associated with vacuolization of the endoplasmic reticulum [69]. An important breakthrough resulted from the demonstration that phalloidin could induce the formation of microfilament-like filamentous structures both in vivo and in preparations of cytoplasmic membrane fragments [70]. Importantly, a nontoxic derivative of phalloidin, desmethylphalloinsulfoxide, did not induce these structures [70]. Two years later, following the lead of Spudich and Lin [64], it was shown that phalloidin could drive the polymerization of monomeric G-actin from rabbit muscle [71]. This polymerization could be inhibited by preincubation with cytochalasin, and the phalloidin-induced filaments produced both in vitro and in vivo could, like conventional actin filaments, be decorated with heavy meromyosin. Thus phalloidin, like cytochalasin, played an important role in connecting the morphologically defined microfilaments with the muscle actin protein.

Advances in Actin Biology Driven by Chemical Inhibitors

The number of literature citations of phalloidin or cytochalasin has increased steadily since the discovery of their mechanism of action (Figure 3). Although the general cell impermeability of phalloidin has limited its use in live cells, an important development has been the utilization of flourescently labeled phalloidin for staining filamentous actin in fixed tissues and cultured cells ([72]; Figure 4). Jasplakinolide, a natural product isolated from a marine sponge, also stabilizes actin polymers yet is cell permeable and has been used in live cells to investigate the importance of filament disassembly in, for example, lamellipodial extension [73]. The cytochalasins rapidly enter living cells, disrupt the actin cytoskeleton (see Figure 4), and have been used to implicate this structure in a various processes. Indeed, as Carter observed in his original description of the activity of cytochalasin:

> By interfering with specific cell activities such as cytoplasmic cleavage and cell movement, they should prove useful as research tools for investigating these important aspects of cell biology. [44]

The identification of the latrunculin family of actin monomer binding drugs deserves mention for its particular contribution to the study of the actin cytoskeleton of yeast. Originally identified as a toxic agent in the marine sponge *Latrunculia magnifica*, the latrunculins (A and B) were shown to disrupt the actin cytoskeleton in mammalian tissue culture cells [74]. Based on these initial observations, latrunculin was shortly thereafter shown to interact directly with monomeric actin in a 1:1 complex, preventing its incorporation into filaments [75].

Although actin is an essential gene in *Saccharomyces cerevisiae* [76], temperature-sensitive mutants in the actin gene have allowed some phenotypic analysis of actin mutations on an approximately 1 hr time scale [77]. The use of latrunculin in *S. cerevisiae*, however, allowed a first look at the acute phase of actin perturbation. Within minutes of addition, latrunculin caused the loss of filamentous actin in a reversible manner [78]. Cytochalasin, by contrast, had no effect, most likely due to cell permeability issues [79]. Using latrunculin, then, allowed the authors to demonstrate that even in the nonmotile yeast cell, the actin cytoskeleton exhibits dynamic de- and repolymerization like in mammalian cells, suggesting that dynamicity of the actin cytoskeleton is a universal phenomenon.

The rapid onset and effectiveness of latrunculin in yeast has led to its continuing wide use in exploring the role of the actin cytoskeleton of this organism in such areas as cell polarity [80], spindle orientation [81], and endocytosis [82]. Indeed, the widespread use of latrunculin when temperature-sensitive mutations also exist is a testament to the advantages offered by small molecule inhibitors (see below). Latrunculin has also become more widely used in mammalian cells. The observation that latrunculin and cytochalasin produce distinct effects in mammalian cells (Figure 4; [83]) is indicative of different mechanisms of action. Whereas cytochalasin binds both monomeric actin and filament 'barbed' ends, latrunculin binds exclusively to actin monomers, making the interpretation of latrunculin experiments somewhat more straightforward [75, 84]. Furthermore, cytochalasin B (but not cytochalasin D) was shown to also inhibit glucose uptake by cells [85, 86] raising questions of specificity, whereas the identification of mutants in actin that confer latrunculin resistance in *S. cerevisiae* have strongly suggested that the interaction of latrunculin and actin in yeast is highly specific [78]. Cytochalasin and latrunculin are each, therefore, unique probes of actin function offering distinct mechanisms of perturbation.

Small Molecules as Tools in Crystallography

Crystallographic analysis of cytoskeletal proteins is complicated by the tendency of individual subunits to polymerize into filaments at high concentrations. This

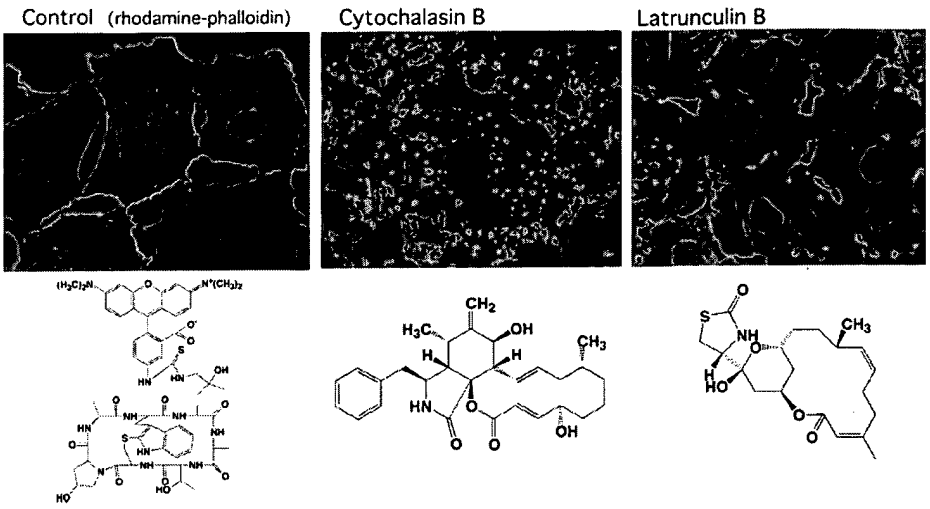


Figure 4. Structure and Effect of Classical Actin-Perturbing Agents

B-SC-1 cells treated with 10.4 μM cytochalasin B, 200 nM latrunculin B, or vehicle control were fixed and stained with rhodamine-conjugated phalloidin.

difficulty has been overcome in the case of actin by chemical modification of the actin monomer [87], by cocrystallizing actin with monomer binding proteins that prevent polymerization [88, 89, 90, 91], and by the use of the small molecule latrunculin [84]. Using the converse approach, Nogales et al. [92] used taxol and zinc to stabilize sheets of tubulin in order to derive an atomic model of the αβ tubulin dimer. Thus, small molecules have also been important tools for investigating the molecular architecture of actin and tubulin at the atomic level.

In and Out in the Blink of an Eye

Chemical inhibitors have illuminated cytoskeletal function not only through the study of compound-treated cells. Important observations have been made of cells during the "washout" or recovery period when compound-treated cells are washed into media lacking the inhibitor. Indeed, the rapid reversibility of colchicine and nocodzole was instrumental in revealing the microtubule-nucleating role of the microtubule organizing center (MTOC). Upon washout of cells treated with nocodazole or colcemid, microtubules are observed to preferentially regrow from this perinuclear structure, suggesting that the MTOC plays a normal role in the nucleation of new microtubules [93, 94, 95]. These observations were confirmed using cold-induced depolymerization of microtubules followed by rewarming [96].

Paul Forscher used a similar approach using cytochalasin to investigate actin dynamics in the neuronal growth cone [97]. Time-lapse video images revealed that on addition of cytochalasin, the actin cytoskeletal matrix within the growth cone disappeared by first receding away from the plasma membrane as an intact unit at a rate of 3–6 μm/min (see Figure 5). On cytochalasin washout, the matrix reappeared first at the plasma membrane and widened and extended toward the rear of the growth cone at an identical 3–6 μm/min. These observations strongly suggested that new actin assembly occurs proximal to the plasma membrane and that the entire actin network of the growth cone translocates centripetally back toward the axon. Importantly, this polymerization and retrograde actin flow from the leading edge is now thought to drive protrusion during cell migration [98].

New Screens, New Molecules

The majority of the most widely used cytoskeletal inhibitors today are natural products that initially drew interest for their toxicity or potential medicinal utility. The collective human experience can thus be seen as a rather "low-throughput" screen for bioactive molecules in the natural world. Recent advances in combinatorial chemistry and high-throughput bioassay screening, however, promise to rapidly increase the speed and efficiency of this process and allow it to be directed toward the identification of small molecules with a particular activity. The actin and tubulin networks of cells consist not only of the filaments and tubules themselves. A large number of regulatory and structural proteins, including motors, crosslinkers, depolymerizers, and filament bundlers, can act to create and organize these assemblies. Several researchers have begun to conduct screens for small molecules targeting components other than actin and tubulin themselves. New inhibitors and a few of these new screens are presented below and are intended to illustrate the diversity of target classes and approaches.

Motor Protein Inhibitors

The recent identification of the Eg5 kinesin inhibitor monastrol demonstrates the success of using whole-cell-based assays to identify inhibitors of the cytoskeleton. Mayer et al. [99] screened for compounds that would induce mitotic arrest in tissue culture cells as assayed by an antibody to the mitotic form of the nucleolin protein. Since compounds that perturb microtubule dynamics (e.g., nocodazole and taxol) can cause mitotic

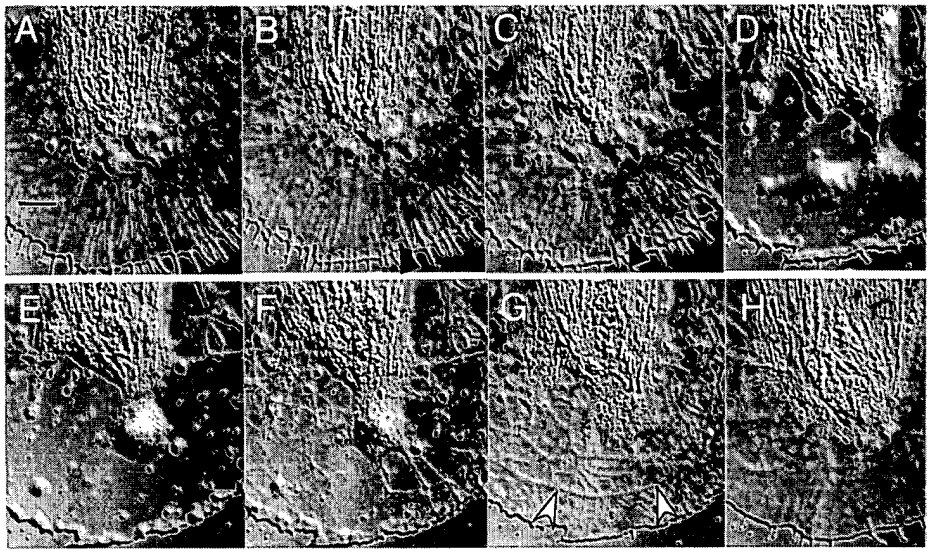


Figure 5. Demonstration that the Actin Network Exhibits Retrograde Flow from the Plasma Membrane, Where New Actin Polymerizes, Centripetally toward the Cell Body in *Aplysia* Neural Growth Cones

Reproduced from The Journal of Cell Biology, 1988, *107*, 1505–1516 by copyright permission of The Rockefeller University Press. Time-lapse differential interference contrast video images of a growth cone after treatment with 10 μM cytochalasin. (A)–(E) represent 0, 0.5, 1, 3, and 9 min of treatment, respectively. The actin network is seen to recede from the plasma membrane at 3–6 μm/min (arrowheads in B and C). Cytochalasin was then removed from the culture media, and cells were allowed to recover for 1, 3, and 17 min (F–H, respectively). The border of the new, advancing actin network (empty arrowheads in G) also migrates at 3–6 μm/min toward the axon.

arrest, active compounds were subsequently counterscreened to eliminate molecules that directly affect microtubule dynamics in vitro. One of the resulting compounds produced a remarkable reorganization of the mitotic spindle of treated cells. Instead of bipolar spindles, monastrol-treated cells produced monoastral spindles. A similar phenotype had been previously observed both in vitro [100] and in vivo [101] on inhibition of the mitotic kinesin Eg5 using anti-Eg5 antibodies. Indeed, in vitro experiments showed that monastrol inhibited microtubule motility powered by Eg5, whereas a structurally related compound that did not cause monoasters in cells did not [99].

Importantly, monastrol appears to show remarkable specificity despite its low micromolar IC_{50} against Eg5. Microtubule arrays in interphase cells appear to be completely unaffected, and effects of monastrol are rapidly reversed on washout of the compound. The specificity, reversibility, and cell permeability of monastrol promises that this compound will be an invaluable tool to help reveal the functions of Eg5 during mitosis. Indeed, using monastrol it was shown that the motor activity of Eg5 was not required for its normal spindle localization [102]. In addition, perturbation of spindle function by monastrol allowed Kapoor et al. [103] to probe the spindle-assembly checkpoint without directly affecting microtubule dynamics. An Eg5 inhibitor with nanomolar affinity was recently reported [104], and this or related compounds will be tested as anticancer drugs in humans.

Recently, two new cell-permeable myosin motor inhibitors were identified in pure protein screens for inhibitors of skeletal muscle myosin II [105] and nonmuscle myosin II [106]. The myosin superfamily of actin-based motors is large and diverse [107], and small molecules that can discriminate between members will allow detailed study of their unique functions in vivo.

A conceptually different approach to chemical inhibition of motor proteins has now been used with both kinesin and myosin motors. Pioneered by Shokat [108], this method involves expression of mutated nucleotide binding proteins with engineered sensitivity to a nucleotide analog. The first proof of principle of this approach for motor proteins involved a single amino acid mutation in the nucleotide binding pocket of kinesin [109]. This mutation conferred sensitivity to the nonhydrolyzeable ATP analog cyclopentyl-adenylyliminodiphosphate, which does not inhibit the wild-type protein. Thus, these authors demonstrated a new experimental approach for the specific inhibition of motor proteins. Holt et al. [110] have recently utilized this approach to address the function of the myosin-1c protein in adaptation of the hair cells in the sensory epithelium of the inner ear. A mutation of the myosin-1c nucleotide binding pocket was generated that would accommodate an N^6-modified ADP analog but that would not prevent its utilization of ATP [111]. Sensory epithelia isolated from mice expressing the mutant protein behaved normally electrophysiologically. When exposed to the ADP analog, however, a loss of the adaptive response to hair cell deflection was observed, demonstrating a crucial role for myosin-1c in hair cell adaptation.

Signaling Protein Inhibitors

Using an approach intermediate between cell-based screens and pure protein assays, Peterson et al. [112] screened for small molecules that would inhibit a signal-

ing pathway controlling the nucleation and polymerization of actin filaments in a cytoplasmic extract. By screening for inhibitors of an entire pathway, these authors screened multiple potential targets, both known and unknown, allowing the biology to dictate the best targets. Interestingly, two inhibitors of a signal integrating protein, the neural Wiskott Aldrich Syndrome protein (N-WASP), were identified using this screen ([112]; J.R.P., L. Bickford, A. Kim, M. Kirschner, and M. Rosen, unpublished data). N-WASP exists in an autoinhibited state that can be activated by binding to signaling molecules such as active cdc42, Nck, or phosphatidylinositol 4,5-bisphosphate [113]. On binding its activators, N-WASP undergoes a conformational change that allows it to activate the Arp2/3 complex, an actin nucleating protein complex [113]. Both of the inhibitors, 187-1, a 14 amino acid cyclic peptide, and wiskostatin, an N-alkylated carbazole derivative, appear to inhibit N-WASP by stabilizing the autoinhibited conformation, thus preventing subsequent activation of the Arp2/3 complex (Figure 6). The identification of two chemically distinct inhibitors of N-WASP suggests that this protein is not only an important signaling node but potentially also an important locus for inhibitors of this pathway.

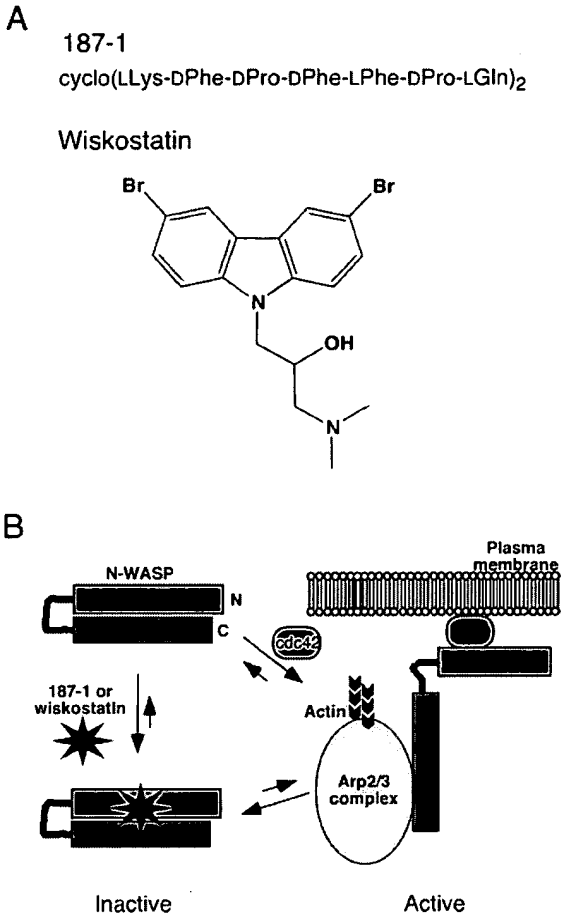


Figure 6. Inhibitors of N-WASP Block Actin Filament Assembly by Stabilizing the Autoinhibited Conformation of N-WASP

(A) Structures of two N-WASP inhibitors, the 14 amino acid cyclic peptide 187-1 and wiskostatin.

(B) Signaling molecules including Cdc42 bind N-WASP, relieving its intrinsic autoinhibition and exposing a C-terminal domain that can bind and activate the Arp2/3 complex to nucleate a new actin filament. 187-1 and wiskostatin attenuate this signaling cascade by stabilizing the autoinhibited conformation of N-WASP, antagonizing activation by upstream signaling molecules (based on [112] and J.R.P., L. Bickford, A. Kim, M. Kirschner, and M. Rosen, unpublished data).

Conclusions

Why have chemical approaches to study the cytoskeleton been so successful? One answer must certainly be the swift action of small molecules. Cytoskeletal rearrangements typically occur over seconds, a time scale inaccessible to traditional genetic approaches but addressable by the rapid diffusion of small molecules. This avoids the complications of cellular adaptation/compensation that can arise when using genetic knockout approaches. An alternative to the knockout of genes of interest is the use of temperature-sensitive mutants. This approach sometimes allows relatively rapid protein inactivation and also allows the study of the loss of function of essential genes. Temperature shifts alone, however, have the potential for nonspecific perturbation [114]. Indeed, transcriptional profiling of *Saccharomyces cerevisiae* has demonstrated "massive and rapid alterations in genomic expression" of wild-type strains on temperature shift from 25° to 37°C [115, 116]. The quick reversibility of many inhibitors allows acute temporal control over the inhibition as well as investigation of the 'recovery phase.' Additionally, the apparent target specificity shown by inhibitors like nocodazole and latrunculin allows almost genetic knockout-like inactivation of individual components of the complex cytoskeletal network. Finally, the functional roles of actin and tubulin appear to be broadly conserved across the eukaryotes. Therefore, the study of different experimental systems is greatly benefited by reagents that are neither species- nor cell-type dependent.

More speculatively, perhaps the cytoskeleton has been well served with small molecule inhibitors simply because many of its components are "druggable." A remarkable number of small molecules have been identified that directly target microtubules [40]. In one unbiased screen for compounds that would induce mitotic arrest, 38% of the initial 'hits' proved to directly affect microtubule polymerization and/or stability [99, 117]. These results raise an important question: what makes tubulin such a good drug target? Further study of the structural basis for the binding and mechanism of action of these inhibitors should help shed light on this issue. In addition, the mitotic checkpoint may be sensitive to and amplify even subtle perturbations of the spindle.

Intriguingly, of the small molecule targets discussed here (actin, tubulin, Eg5, muscle myosin, nonmuscle myosin, N-WASP), all are proteins that undergo reversible conformational changes as part of their functional cycles. Indeed, several of their inhibitors (latrunculin, 187-1, wiskostatin) appear to act by blocking these conformational changes, suggesting that target "inhibitability" and conformational flexibility may be related [84, 112]. In this context, it is interesting to note the lack of

small molecules that target the third major cytoskeletal system, the 'intermediate' filaments. These structures are thought to play predominantly a more rigid, structural role, and their inherent stability, then, may be less amenable to disruption by small molecules.

The newly discovered inhibitors discussed above represent only the tip of the iceberg of molecules yet to be identified using high-throughput screening technology. The evolving technology coupling combinatorially synthesized compound libraries with cytoskeleton-oriented screens, whether pure protein, extract-, or cell-based, promises to rapidly deliver to chemical genetics the equivalent of "saturation mutagenesis:" all cytoskeletal targets screened versus a vast universe of small molecules. The greatest challenge to the chemical approach remains the issue of specificity, ensuring that the phenotype caused by a compound is indeed due the inhibition of only its supposed "target." Confirmation using independent approaches will help address this concern. Yet nature has provided us with remarkable examples of small molecules that appear to act on particular cytoskeletal proteins with exquisite specificity. This has been demonstrated directly by the fact that mutations conferring resistance to taxol and latrunculin can be identified in the tubulin and actin genes, respectively [118, 77]. These encouraging examples suggest that if we seek, we shall find.

Acknowledgments

We gratefully acknowledge Bill Breiher, Aaron Straight, and Marc Kirschner for critical reading of the manuscript.

References

1. Cohen, A., Cook, J., and Roe, E. (1940). Colchicine and related compounds. J. Chem. Soc. (London) 194–197.
2. Eigsti, O.J., and Dustin, P. (1955). Colchicine in Agriculture, Medicine, Biology, and Chemistry. (Ames, IA: Iowa State College Press).
3. Dustin, P. (1978). Microtubules. (Berlin-Heidelberg: Springer-Verlag).
4. Tjio, J.H., and Levan, A. (1956). The chromosome number of man. Hereditas *42*, 1–6.
5. Taylor, E.W. (1965). The mechanism of colchicine inhibition of mitosis. J. Cell Biol. *25*, 145–160.
6. Borisy, G.G., and Taylor, E.W. (1967). The mechanism of action of colchicine. Binding of colchincine-^{3}H to cellular protein. J. Cell Biol. *34*, 525–533.
7. Borisy, G.G., and Taylor, E.W. (1967). The mechanism of action of colchicine. Colchicine binding to sea urchin eggs and the mitotic apparatus. J. Cell Biol. *34*, 535–548.
8. Shelanski, M.L., and Taylor, E.W. (1967). Isolation of a protein subunit from microtubules. J. Cell Biol. *34*, 549–554.
9. Weisenberg, R.C., Borisy, G.G., and Taylor, E.W. (1968). The colchicine-binding protein of mammalian brain and its relation to microtubules. Biochemistry *7*, 4466–4479.
10. Mohri, H. (1968). Amino-acid composition of "Tubulin" constituting microtubules of sperm flagella. Nature *217*, 1053–1054.
11. Wani, M.C., Taylor, H.L., Wall, M.E., Coggon, P., and McPhail, A.T. (1971). Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from Taxus brevifolia. J. Am. Chem. Soc. *93*, 2325–2327.
12. Rowinsky, E.K., and Donehower, R.C. (1991). Taxol: twenty years later, the story unfolds. J. Natl. Cancer Inst. *83*, 1778–1781.
13. Wall, M.E., and Wani, M.C. (1995). Camptothecin and taxol: discovery to clinic—thirteenth Bruce F. Cain Memorial Award Lecture. Cancer Res. *55*, 753–760.
14. Fuchs, D.A., and Johnson, R.K. (1978). Cytologic evidence that taxol, an antineoplastic agent from Taxus brevifolia, acts as a mitotic spindle poison. Cancer Treat. Rep. *62*, 1219–1222.
15. Schiff, P.B., Fant, J., and Horwitz, S.B. (1979). Promotion of microtubule assembly in vitro by taxol. Nature *277*, 665–667.
16. Heidemann, S.R., and Gallas, P.T. (1980). The effect of taxol on living eggs of *Xenopus laevis*. Dev. Biol. *80*, 489–494.
17. Schiff, P.B., and Horwitz, S.B. (1980). Taxol stabilizes microtubules in mouse fibroblast cells. Proc. Natl. Acad. Sci. USA *77*, 1561–1565.
18. Vallee, R.B. (1982). A taxol-dependent procedure for the isolation of microtubules and microtubule-associated proteins (MAPs). J. Cell Biol. *92*, 435–442.
19. Vale, R.D., Reese, T.S., and Sheetz, M.P. (1985). Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. Cell *42*, 39–50.
20. De Brabander, M., Van de Veire, R., Aerts, F., Geuens, G., Borgers, M., Desplenter, L., and De Cree, J. (1975). In Microtubules and Microtubule Inhibitors, M. Borgers and M. De Brabander, eds. (Amsterdam: North-Holland Publishing Company), pp. 509–521.
21. Hoebeke, J., Van Nijen, G., and De Brabander, M. (1976). Interaction of oncodazole (R 17934), a new anti-tumoral drug, with rat brain tubulin. Biochem. Biophys. Res. Commun. *69*, 319–324.
22. Thyberg, J., and Moskalewski, S. (1985). Microtubules and the organization of the Golgi complex. Exp. Cell Res. *159*, 1–16.
23. Mareel, M.M., and De Mets, M. (1984). Effect of microtubule inhibitors on invasion and on related activities of tumor cells. Int. Rev. Cytol. *90*, 125–168.
24. Malawista, S.E. (1965). On the action of colchicine; the melanocyte model. J. Exp. Med. *122*, 361–384.
25. Malawista, S.E. (1975). Microtubules and the movement of melanin granules in frog dermal melanocytes. Ann. N Y Acad. Sci. *253*, 702–710.
26. Murphy, D.B. (1975). The mechanism of microtubule-dependent movement of pigment granules in teleost chromatophores. Ann. N Y Acad. Sci. *253*, 692–701.
27. Kung, A.L., Sherwood, S.W., and Schimke, R.T. (1990). Cell line-specific differences in the control of cell cycle progression in the absence of mitosis. Proc. Natl. Acad. Sci. USA *87*, 9553–9557.
28. Murray, A.W., and Kirschner, M.W. (1989). Dominoes and clocks: the union of two views of the cell cycle. Science *246*, 614–621.
29. Rieder, C.L., and Palazzo, R.E. (1992). Colcemid and the mitotic cycle. J. Cell Sci. *102*, 387–392.
30. Hoyt, M.A., Totis, L., and Roberts, B.T. (1991). *S. cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. Cell *66*, 507–517.
31. Li, R., and Murray, A.W. (1991). Feedback control of mitosis in budding yeast. Cell *9*, 519–531.
32. Noble, R.L. (1990). The discovery of the *Vinca* alkaloids–chemotherapeutic agents against cancer. Biochem. Cell Biol. *68*, 1344–1351.
33. Noble, R.L. (1958). Role chance observations in chemotherapy: *Vinca rosea*. Ann. N Y Acad. Sci. *76*, 882–894.
34. Rowinsky, E.K., and Donehower, R.C. (1991). The clinical pharmacology and use of antimicrotubule agents in cancer chemotherapeutics. Pharmacol. Ther. *52*, 35–84.
35. Jordan, M.A., and Wilson, L. (1998). Microtubules and actin filaments: dynamic targets for cancer chemotherapy. Curr. Opin. Cell Biol. *10*, 123–130.
36. Nicolaou, K.C., Roschangar, F., and Vourloumis, D. (1998). Chemical biology of epothilones. Angew. Chem. Int. Ed. Engl. *37*, 2014–2045.
37. Altmann, K.-H. (2001). Microtubule-stabilizing agents: a growing class of important anticancer drugs. Curr. Opin. Chem. Biol. *5*, 424–431.
38. Nicolaou, K.C., Scarpelli, R., Bollbuck, B., Werschkun, B., Pereira, M.M.A., Wartmann, M., Altmann, K.-H., Zaharevitz, D., Gussio, R., and Giannakakou, P. (2000). Chemical synthesis

and biological properties of pyridine epothilones. Chem. Biol. *7*, 593–599.
39. Nicolaou, K.C., Ritzen, A., and Namoto, K. (2001). Recent developments in the chemistry, biology and medicine of the epothilones. Chem. Commun. *17*, 1523–1535.
40. Jordan, A., Hadfield, J.A., Lawrence, N.J., and McGown, A.T. (1998). Tubulin as a target for anticancer drugs: agents which interact with the mitotic spindle. Med. Res. Rev. *18*, 259–296.
41. Kingston, D.G.I. (2000). Recent advances in the chemistry of taxol. J. Nat. Prod. *63*, 726–734.
42. Aldridge, D.C., Armstrong, J.J., Speake, R.N., and Turner, W.B. (1967). The cytochalasins, a new class of biologically active mould metabolites. Chem. Commun. *1*, 26–27.
43. Rothweiler, W., and Tamm, C. (1966). Isolation and structure of phomin. Experientia *22*, 750–752.
44. Carter, S.B. (1967). Effects of cytochalasins on mammalian cells. Nature *213*, 261–264.
45. Dixon, B. (1966). The cytocholasins and cell mechanics. New Sci. *32*, 170–171.
46. Carter, S.B. (1967). Haptotaxis and the mechanism of cell motility. Nature *213*, 256–260.
47. Wolpert, L. (1960). The mechanics and mechanism of cleavage. Int. Rev. Cytol. *10*, 163–216.
48. Marsland, D., and Landau, J.V. (1954). The mechanisms of cytokinesis: temperature-pressure studies on the cortical gel system in various marine eggs. J. Exp. Zool. *125*, 507–539.
49. Mitchison, J.M. (1952). Cell membranes and cell division. Symp. Soc. Exp. Biol. *6*, 105–127.
50. Gray, J. (1924). The mechanism of cell-division I. The forces which control the form and cleavage of the eggs of *Echinus esculentus*. Proc. Cambridge Phil. Soc. Biol. Sci. *1*, 164–188.
51. Dan, J.C. (1948). On the mechanism of astral cleavage. Physiol. Zool. *21*, 191–218.
52. Schroeder, T.E. (1970). The contractile ring. I. Fine structure of dividing mammalian (HeLa) cells and the effects of cytochalasin B. Z. Zellforsch. Mikrosk. Anat. *109*, 431–449.
53. Schroeder, T.E. (1969). The role of "contractile ring" filaments in dividing *Arbacia* egg. Biol. Bull. *137*, 413–414.
54. Buckley, I.K., and Porter, K.R. (1967). Cytoplasmic fibrils in living cultured cells. Protoplasma *64*, 349–380.
55. Ishikawa, H., Bischoff, R., and Holtzer, H. (1969). Formation of arrowhead complexes with heavy meromyosin in a variety of cell types. J. Cell Biol. *43*, 312–328.
56. Pollard, T.D., Shelton, E., Weihing, R.R., and Korn, E.D. (1970). Ultrastructural characterization of F-actin isolated from *Acanthamoeba castellanii* and identification of cytoplasmic filaments as F-actin by reaction with rabbit heavy meromyosin. J. Mol. Biol. *50*, 91–97.
57. Tilney, L.G., and Mooseker, M. (1971). Actin in the brush-border of epithelial cells of the chicken intestine. Proc. Natl. Acad. Sci. USA *68*, 2611–2615.
58. Woolley, D.E. (1970). An actin-like protein from amoebae of *Dictyostelium discoideum*. Fed. Proc. *29*, 667.
59. Nachmias, V.T., Huxley, H.E., and Kessler, D. (1970). Electron microscope observations on actomyosin and actin preparations from *Physarum polycephalum*, and on their interaction with heavy meromyosin subfragment I from muscle myosin. J. Mol. Biol. *50*, 83–90.
60. Pollard, T.D., and Korn, E.D. (1971). Filaments of *Amoeba proteus*. II. Binding of heavy meromyosin by thin filaments in motile cytoplasmic extracts. J. Cell Biol. *48*, 216–219.
61. Yamada, K.M., Spooner, B.S., and Wessells, N.K. (1970). Axon growth: roles of microfilaments and microtubules. Proc. Natl. Acad. Sci. USA *66*, 1206–1212.
62. Spooner, B.S., and Wessells, N.K. (1970). Effects of cytochalasin B upon microfilaments involved in morphogenesis of salivary epithelium. Proc. Natl. Acad. Sci. USA *66*, 360–364.
63. Wrenn, J.T., and Wessells, N.K. (1970). Cytochalasin B: effects upon microfilaments involved in morphogenesis of estrogen-induced glands of oviduct. Proc. Natl. Acad. Sci. USA *66*, 904–908.
64. Spudich, J., and Lin, S. (1972). Cytochalasin B, its interaction with actin and actomyosin from muscle. Proc. Natl. Acad. Sci. USA *69*, 442–446.
65. Lynen, F., and Wieland, U. (1937). Uber die giftstoffe des Knollenblätterpilzes IV. Justus Liebigs Ann. Chem. *533*, 93–117.
66. Wieland, T., and Wieland, O. (1959). Chemistry and toxicology of the toxins of *Amanita phalloides*. Pharmacol Rev. *11*, 87–107.
67. Raab, H.A. (1932). Beitrage zur Kenntnis des Giftstoffs der Knollenblätterpilze. Hoppe-Seyler's Z. Physiol. Chem. *207*, 157–181.
68. Wieland, T. (1983). The toxic peptides from *Amanita* mushrooms. Int. J. Pept. Protein Res. *22*, 257–276.
69. Miller, F., and Wieland, O. (1967). Elektronenmikroskopische Untersuchungen der Leber von Maus und Ratte bei acuter Phalloidin-Vergiftung. Virchows Arch. A Pathol. Anat. Histol. *343*, 83–99.
70. Govindan, V.M., Faulstich, H., Wieland, T., Agostini, B., and Hasselbach, W. (1972). In-vitro effect of phalloidin on a plasma membrane preparation from rat liver. Naturwissenschaften *59*, 521–522.
71. Lengsfeld, A.M., Low, I., Wieland, T., Dancker, P., and Hasselbach, W. (1974). Interaction of phalloidin with actin. Proc. Natl. Acad. Sci. USA *71*, 2803–2807.
72. Wulf, E., Deboben, A., Bautz, F.A., Faulstich, H., and Wieland, T. (1979). Flouorescent phallotoxin, a tool for the visualization of cellular actin. Proc. Natl. Acad. Sci. USA *76*, 4498–4502.
73. Cramer, L.P. (1999). Role of actin-filament disassembly in lamellipodium protrusion in motile cells revealed using the drug jasplakinolide. Curr. Biol. *9*, 1095–1105.
74. Spector, I., Shochet, N.R., Kashman, Y., and Groweiss, A. (1983). Latrunculins: novel marine toxins that disrupt microfilament organization in cultured cells. Science *219*, 493–495.
75. Coue, M., Brenner, S.L., Spector, I., and Korn, E.D. (1987). Inhibition of actin polymerization by latrunculin A. FEBS Lett. *213*, 316–318.
76. Shortle, D., Haber, J.E., and Botstein, D. (1982). Lethal disruption of the yeast actin gene by integrative DNA transformation. Science *217*, 371–373.
77. Novick, P., and Botstein, D. (1985). Phenotypic analysis of temperature-sensitive yeast actin mutants. Cell *40*, 405–416.
78. Ayscough, K.R., Stryker, J., Pokala, N., Sanders, M., Crews, P., and Drubin, D.G. (1997). High rates of actin filament turnover in budding yeast and roles for actin in establishment and maintenance of cell polarity revealed using the actin inhibitor latrunculin-A. J. Cell Biol. *137*, 399–416.
79. Greer, C., and Schekman, R. (1982). Actin from *Saccharomyces cerevisiae*. Mol. Cell. Biol. *2*, 1270–1278.
80. Karpova, T.S., Reck-Peterson, S.L., Elkind, N.B., Mooseker, M.S., Novick, P.J., and Cooper, J.A. (2000). Role of actin and Myo2p in polarized secretion and growth of Saccharomyces cerevisiae. Mol. Biol. Cell *11*, 1727–1737.
81. Theesfeld, C.L., Irazoqui, J.E., Bloom, K., and Lew, D.J. (1999). The role of actin in spindle orientation changes during the *Saccharomyces cerevisiae* cell cycle. J. Cell Biol. *146*, 1019–1032.
82. Ayscough, K.R. (2000). Endocytosis and the development of cell polarity in yeast require a dynamic F-actin cytoskeleton. Curr. Biol. *10*, 1587–1590.
83. Spector, I., Shochet, N.R., Blasberger, D., and Kashman, Y. (1989). Latrunculins–novel marine macrolides that disrupt microfilament organization and affect cell growth: I. Comparison with cytochalasin D. Cell Motil. Cytoskeleton *13*, 127–144.
84. Morton, W.M., Ayscough, K.R., and McLaughlin, P.J. (2000). Latrunculin alters the actin-monomer subunit interface to prevent polymerization. Nat. Cell Biol. *2*, 376–378.
85. Zigmond, S.H., and Hirsch, J.G. (1972). Cytochalasin B: inhibition of D-2-deoxyglucose transport into leukocytes and fibroblasts. Science *176*, 1432–1434.
86. McDaniel, M., Roth, C., Fink, J., Fyfe, G., and Lacy, P. (1975). Effects of cytochalasins B and D on alloxan inhibition of insulin release. Biochem. Biophys. Res. Commun. *66*, 1089–1096.
87. Otterbein, L.R., Graceffa, P., and Dominguez, R. (2001). The crystal structure of uncomplexed actin in the ADP state. Science *293*, 708–711.
88. Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., and Holmes,

K.C. (1990). Atomic structure of the actin:DNase I complex. Nature *347*, 37–44.

89. McLaughlin, P.J., Gooch, J.T., Mannherz, H.G., and Weeds, A.G. (1993). Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. Nature *364*, 685–692.
90. Robinson, R.C., Mejillano, M., Le, V.P., Burtnick, L.D., Yin, H.L., and Choe, S. (1999). Domain movement in gelsolin: a calcium-activated switch. Science *286*, 1939–1942.
91. Schutt, C.E., Myslik, J.C., Rozycki, M.D., Goonesekere, N.C., and Lindberg, U. (1993). The structure of crystalline profilin-beta-actin. Nature *365*, 810–816.
92. Nogales, E., Wolf, S.G., and Downing, K.H. (1998). Structure of the alpha beta tubulin dimer by electron crystallography. Nature *391*, 199–203.
93. Brinkley, B.R., Stubblefield, E., and Hsu, T.C. (1967). The effects of colcemid inhibition and reversal on the fine structure of the mitotic apparatus of Chinese hamster cells in vitro. J. Ultrastruct. Res. *19*, 1–18.
94. Osborn, M., and Weber, K. (1976). Cytoplasmic microtubules in tissue culture cells appear to grow from an organizing structure towards the plasma membrane. Proc. Natl. Acad. Sci. USA *73*, 867–871.
95. De Brabander, M., Geuens, G., De Mey, J., and Joniau, M. (1981). Nucleated assembly of mitotic microtubules in living PTK2 cells after release from nocodazole treatment. Cell Motil. *1*, 469–483.
96. Rieder, C.L., and Borisy, G.G. (1981). The attachment of kinetochores to the pro-metaphase spindle in PtK1 cells. Recovery from low temperature treatment. Chromosoma *82*, 693–716.
97. Forscher, P., and Smith, S.J. (1988). Actions of cytochalasins on the organization of actin filaments and microtubules in a neuronal growth cone. J. Cell Biol. *107*, 1505–1516.
98. Welch, M.D., Mallavarapu, A., Rosenblatt, J., and Mitchison, T.J. (1997). Actin dynamics in vivo. Curr. Opin. Cell Biol. *9*, 54–61.
99. Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., and Mitchison, T.J. (1999). Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. Science *286*, 971–974.
100. Sawin, K.E., LeGuellec, K., Philippe, M., and Mitchison, T.J. (1992). Mitotic spindle organization by a plus-end-directed microtubule motor. Nature *359*, 540–543.
101. Blangy, A., Lane, H.A., d'Herin, P., Harper, M., Kress, M., and Nigg, E.A. (1995). Phosphorylation by p34cdc2 regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation in vivo. Cell *83*, 1159–1169.
102. Kapoor, T.M., and Mitchison, T.J. (2001). Eg5 is static in bipolar spindles relative to tubulin: evidence for a static spindle matrix. J. Cell Biol. *154*, 1125–1133.
103. Kapoor, T.M., Mayer, T.U., Coughlin, M.L., and Mitchison, T.J. (2000). Probing spindle assembly mechanisms with monastrol, a small molecule inhibitor of the mitotic kinesin, Eg5. J. Cell Biol. *150*, 975–988.
104. Finer, J.T., Bergnes, G., Fend, B., Smith, W.W., and Chabala, J.C. (2001). Methods and compositions utilizing quinazolinones. WIPO. WO 01/30768 A1.
105. Cheung, A., Dantzig, J.A., Hollingworth, S., Baylor, S.M., Goldman, Y.E., Mitchison, T.J., and Straight, A.F. (2001). A small-molecule inhibitor of skeletal muscle myosin II. Nat. Cell Biol. *4*, 83–88.
106. Cheung, A., Westwood, N.J., Chen, I., Mitchison, T.J., and Straight, A.F. (2001). Blebbistatin: a cell permeable inhibitor of non-muscle myosin II. Mol. Biol. Cell Suppl. *12*, 271a.
107. Sellers, J.R. (2000). Myosins: a diverse superfamily. Biochim. Biophys. Acta *1496*, 3–22.
108. Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., et al. (2000). Unnatural ligands for engineered proteins: new tools for chemical genetics. Annu. Rev. Biophys. Biomol. Struct. *29*, 577–606.
109. Kapoor, T.M., and Mitchison, T.J. (1999). Allele-specific activators and inhibitors for kinesin. Proc. Natl. Acad. Sci. USA *96*, 9106–9111.
110. Holt, J.R., Gillespie, S.K., Provance, D.W., Shah, K., Shokat, K.M., Corey, D.P., Mercer, J.A., and Gillespie, P.G. (2002). A chemical-genetic strategy implicates myosin-1c in adaptation by hair cells. Cell *108*, 371–381.
111. Gillespie, P.G., Gillespie, S.K., Mercer, J.A., Shah, K., and Shokat, K.M. (1999). Engineering of the myosin-I beta nucleotide-binding pocket to create selective sensitivity to N(6)-modified ADP analogs. J. Biol. Chem. *274*, 31373–31381.
112. Peterson, J.R., Lokey, R.S., Mitchison, T.J., and Kirschner, M.W. (2001). A chemical inhibitor of N-WASP reveals a new mechanism for targeting protein interactions. Proc. Natl. Acad. Sci. USA *98*, 10624–10629.
113. Zigmond, S.H. (2000). How WASP regulates actin polymerization. J. Cell Biol. *150*, F117–F119.
114. Shogren-Knaak, M.A., Alaimo, P.J., and Shokat, K.M. (2001). Recent advances in chemical approaches to the study of biological systems. Annu. Rev. Cell Dev. Biol. *17*, 405–433.
115. Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. Mol. Biol. Cell *11*, 4241–4257.
116. Causton, H.C., Ren, B., Koh, S.S., Harbison, C.T., Kanin, E., Jennings, E.G., Lee, T.I., True, H.L., Lander, E.S., and Young, R.A. (2001). Remodeling of yeast genome expression in response to environmental changes. Mol. Biol. Cell *12*, 323–337.
117. Haggarty, S.J., Mayer, T.U., Miyamoto, D.T., Fathi, R., King, R.W., Mitchison, T.J., and Schreiber, S.L. (2000). Dissecting cellular processes using small molecules: identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. Chem. Biol. *7*, 275–286.
118. Cabral, F., Abraham, I., and Gottesman, M.M. (1981). Isolation of a taxol-resistant Chinese hamster ovary cell mutant that has an alteration in alpha-tubulin. Proc. Natl. Acad. Sci. USA *78*, 4388–4391.

Small-molecule-based strategies for controlling gene expression

Carilee Denison and Thomas Kodadek

A central goal in chemical biology is to gain control over biological pathways using small molecules, and the mRNA-synthesizing machinery is a particularly important target. New advances in our understanding of transcriptional regulation suggests strategies to manipulate these pathways using small molecules.

Address: Center for Biomedical Inventions, Departments of Internal Medicine and Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard., Dallas, TX 75235-8573, USA.

Correspondence: Thomas Kodadek
E-mail: kodadek@ryburn.swmed.edu

Chemistry & Biology June 1998, **5**:R129–R145
http://biomednet.com/elecref/10745521005R0145

 ISSN 1074-5521

Introduction

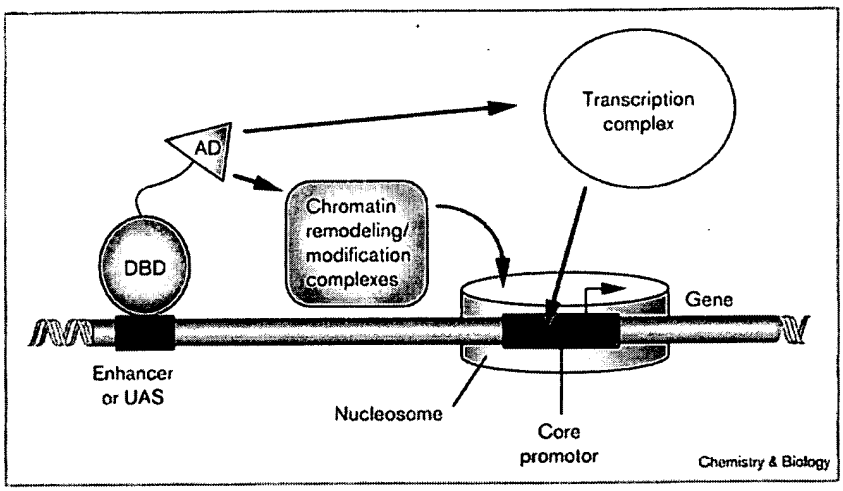
Perhaps the most important regulatory networks in eukaryotic cells are those that control the activity of the mRNA-synthesizing machinery, which includes RNA polymerase II and a host of general transcription factors. This is because the expression of the vast majority of eukaryotic genes is controlled at the level of transcription by activators and repressors. Such proteins function in a gene-specific fashion to up-regulate or down-regulate, respectively, the ability of the transcriptional machinery to synthesize mRNA encoded by the target gene. In turn, the activities of transcriptional activators and repressors are usually regulated by signal transduction cascades that transfer information from the cell membrane to the nucleus. The ultimate goal of molecular medicine is to gain control over these processes and turn particular genes on or off at will. Chemical biology represents one of the two most promising approaches to achieve this goal, the other being gene therapy.

The area of small-molecule-regulated gene expression is reviewed briefly in this article. First, the state of the art of the field is discussed, focusing on work carried out using derivatives of naturally occurring immunosuppressants. Then, the complex protein machinery that mediates mRNA synthesis in eukaryotic cells is discussed, followed by a brief review of the nature of the activators and repressors that regulate the activity of the transcriptional machinery. We then consider potential approaches to finding small molecules that would allow one to manipulate the interactions of native transcription factors and signal-transduction proteins with one another and with nucleic acids. Finally, we consider possibilities for the *de novo* synthesis of small molecules that can regulate transcription directly by acting as activator or repressor mimics.

Natural products as regulators of gene expression

All organisms contain a large number of genes whose levels of expression are controlled by small molecules. For example, many genes involved in biosynthetic pathways are feedback-regulated by a build up of the product of the pathway. Conversely, the expression of catabolic genes, such as those that encode proteins involved in sugar metabolism, are stimulated tremendously when the cell is in a meduim rich in the corresponding sugar. Nature therefore uses small-molecule-controlled gene expression routinely. In most of such cases, binding of the small molecule to a membrane-bound, or sometimes soluble, receptor that is highly specific for recognizing that molecule triggers a cascade that ultimately either stimulates or inhibits the

Figure 1



Schematic view of an activator. The DNA-binding domain (DBD) allows the protein to bind in the vicinity of the target gene. The attached activation domain (AD) makes contacts with chromatin remodeling and modifying complexes as well as the transcription complex, serving to 'open' the chromatin structure and assist association of the transcription complex with the promoter. The activities of the DBD and AD are, to a first approximation, separable and there is not a requirement for a distinct stereochemical relationship between the two domains.

activity of the mRNA-synthesizing machinery at a particular set of genes. In most cases, the ultimate recipients of the signals are transcriptional activators and/or repressors. Activators, which will be discussed in more detail below, generally consist of quasi-separable [1,2] DNA-binding and 'activation' domains (Figure 1). The DNA-binding domain allows the protein to be localized in the vicinity of the target gene, whereas the activation domain binds directly to components of the transcriptional machinery or to chromatin-remodelling complexes, contacts that result in greatly increased transcription of the target genes.

Repressors can block the activity of activators directly, for example by binding to the activation domain and sequestering it from the transcriptional machinery, or they can function indirectly, for example by mediating changes in chromatin structure that block access of transcription proteins to the DNA.

The balance between activators and repressors can be affected by small molecules in a number of ways. For example, the yeast Gal4 protein (Gal4p), a potent activator of genes involved in galactose metabolism [3], is normally muzzled by a specific repressor, Gal80 protein (Gal80p), that binds tightly to the Gal4 activation domain [4]. When the galactose concentration in the medium is increased, the sugar binds to the Gal3 signal-transduction protein, which then acts upon the Gal4–Gal80 complex in an ATP-dependent fashion to relieve the repressive effect of Gal80 and expose Gal4p's activation surface [5]. The result is a huge induction in the expression of the GAL genes.

This theme of small-molecule-dependent release of an activator from a repressive interaction is very common. For example, the important activator NF-κB is sequestered in an inactive complex in the cytoplasm by the protein IκB. Exposure of cells to a variety of stimuli, including small molecules such as phorbol esters, results in the phosphorylation and subsequent proteasome-mediated degradation of IκB. As a result, NF-κB is released and it moves into the nucleus where it binds to DNA and, in concert with other activators, drives the transcription of genes involved in a number of important processes [6], including inflammation and limb development.

Another example of this type of activation mechanism of particular interest to chemical biologists is the action of the clinically used immunosuppressants FK-506 and cyclosporin A (CsA; Figure 2). Although the examples given above were responses to natural stimuli, developmental signals or cellular stress, FK-506 and CsA are produced by soil microorganisms and are not normally present in human cells. These compounds were discovered on the basis of their ability to suppress immune function. Their mechanism of action, which is now understood in considerable detail [7,8], provides an excellent paradigm for the kind of molecular tool for manipulating gene expression that chemical biologists would like to have.

A crucial step in the induction of an immune response is the translocation of the nuclear factor of activated T cells (NFAT) from the cytoplasm to the nucleus of T cells, where it activates the transcription of several genes, including that encoding interleukin-2. The event that triggers nuclear translocation is calcineurin-mediated dephosphorylation of cytoplasmic NFAT (Figure 2). Both FK-506 and CsA interfere with this process by first binding with high affinity to their respective target proteins, FKBP [9] and cyclophilin [10]. Remarkably, the composite surfaces of both the FKBP–FK-506 and cyclophilin–CsA complexes bind calcineurin tightly [11] and inhibit its ability to dephosphorylate NFAT [12,13],

Figure 2

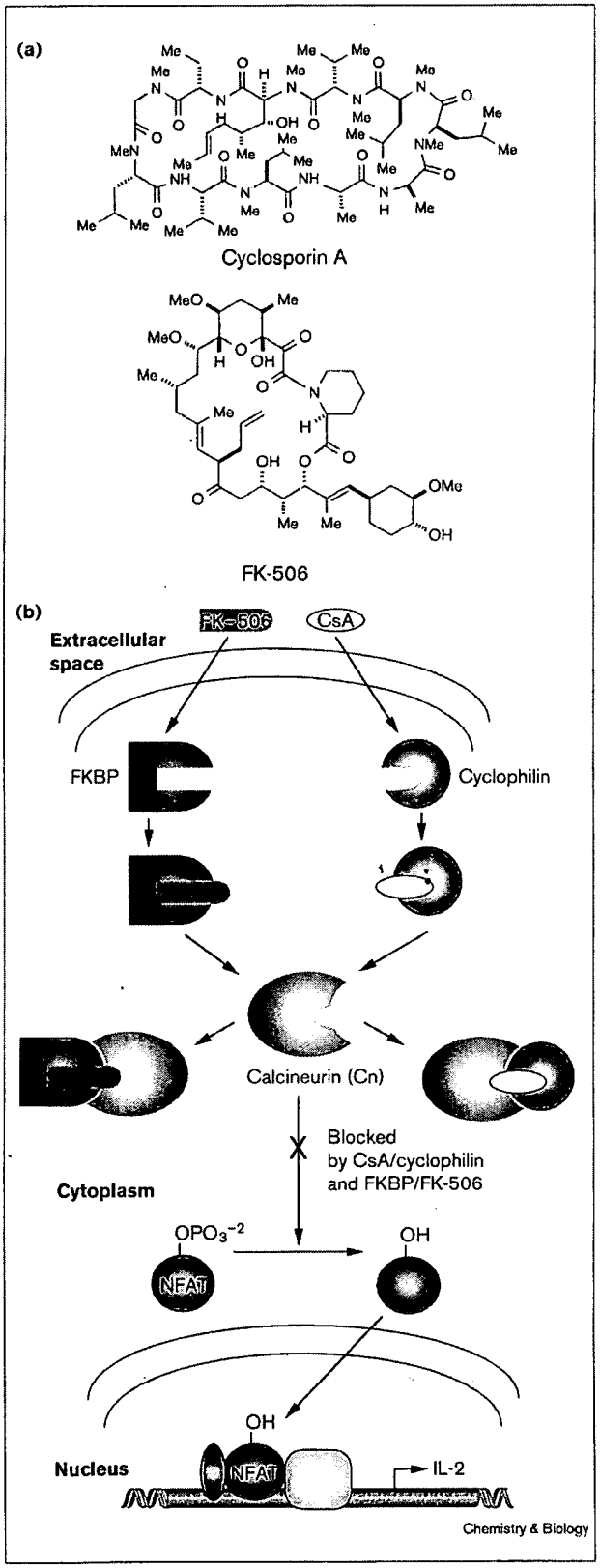


Figure 2

(a) Structures of the immunosuppressants FK-506 and cyclosporin A (CsA). **(b)** Schematic model for the mechanism of action of FK-506 and CsA. Each binds an intracellular receptor (FKBP and cyclophilin, respectively). The protein–small molecule complexes bind and inhibit calcineurin, a phosphatase, which blocks dephosphorylation and nuclear translocation of cytoplasmic nuclear factor for T-cell activation (NFAT). NFAT, along with other gene-specific transcription factors (unlabeled in the figure), is required for transcription of interleukin-2 (IL-2) and other genes involved in generating an immune response.

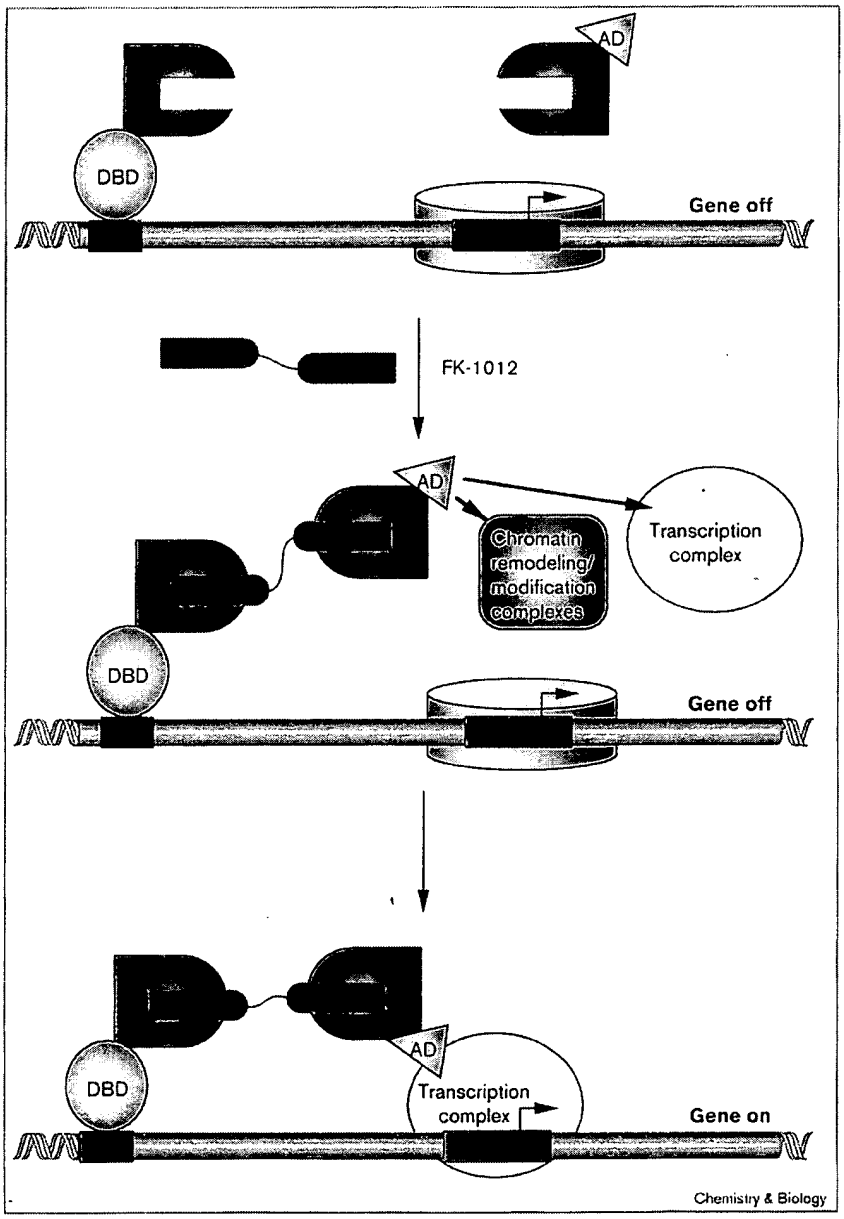
thereby blocking NFAT-activated interleukin-2 gene expression. Thus, CsA and FK-506 can be seen as 'molecular matchmakers' that bring about the association of two proteins that normally do not interact and, in so doing, indirectly down-regulate a particular pathway of gene expression.

Small molecule-mediated control of gene expression in engineered cells

Schreiber, Crabtree and coworkers [14] have used synthetic versions of these remarkable natural products to manipulate protein–protein interactions *in vivo* (Figure 3). The crux of this work is that two proteins of interest are fused to the protein receptors for FK-506 or CsA (FKBP and cyclophilin, respectively) at the DNA level and then expressed in the cell type of interest. The association of the two engineered proteins can then be triggered by the addition of cell-permeable, homodimeric or heterodimeric constructs in which two immunosuppressant molecules have been linked covalently ($(FK\text{-}506)_2$, CsA_2 or FK-CsA) [15]. If mere proximity of the tagged proteins, as opposed to specific interactions, is sufficient to elicit a biological response, then this will occur. For example, an FK-506 homodimer (FK-1012) was introduced into cells engineered to express two fusion proteins, one in which FKBP was linked to the activation domain of a transcription factor (NF3V1) and another in which FKBP was fused to the DNA-binding domain of a transcription factor (GF3 or HF3). As transcriptional activation requires physical linkage of DNA-binding and activation domains (see above), but not any sort of specific interaction between these domains, FK-1012-induced association of these fusion proteins resulted in the transcription of the otherwise silent target genes [16]. Related chemically induced dimerization experiments have been performed using fusions of FKBP or cyclophilin with signal transduction factors, such as the Fas and TCR cell-surface receptors Raf, Src or Sos. These experiments that manipulate signal transduction factors also result in the control of gene expression, but intervene at a point far upstream of the actual transcription machinery.

These experiments illustrate the exciting potential of using small molecules to control gene expression in living cells. They promise to allow scientists and doctors to turn

Figure 3



Manipulation of chimeric transcription factors using a small molecule allows the expression of a reporter gene to be regulated by a cell-permeable small molecule. FK-1012 is a synthetic homodimer containing two tethered FK-506 molecules (see Figure 2a).

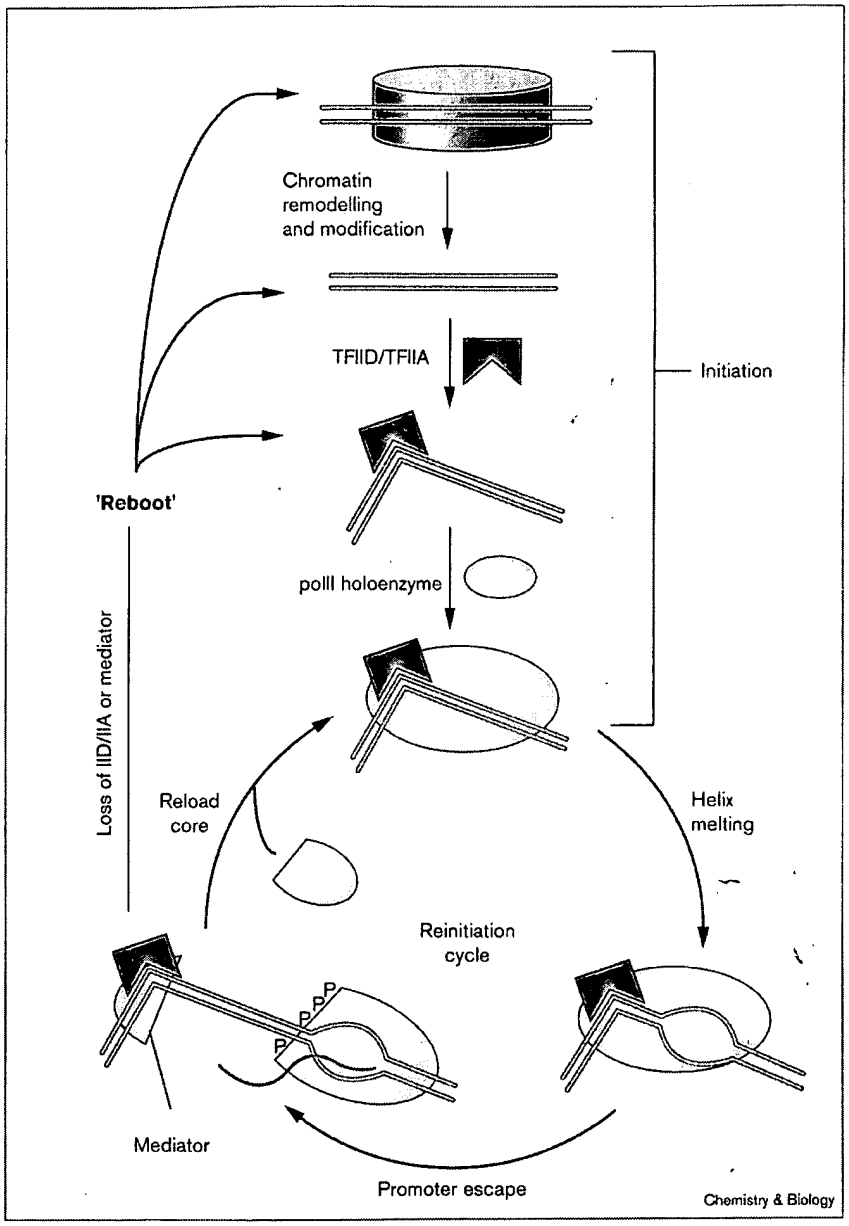
specific genes on and off at will using cell-permeable molecules. As the vast majority of regulation over biological pathways occurs at the level of signal transduction/transcription, such a technology would revolutionize molecular biology and medicine. The limitation with current technology is that immunosuppressant-derived dimerizers can only allow the researcher to manipulate artificial proteins in which an immunophilin or FKBP has been fused to the protein(s) of interest. A key goal is therefore to find compounds that can be used to manipulate the interactions of wild-type macromolecules in nonengineered cells.

The transcription cycle: a complex symphony or a Texas two-step?

To tackle the problem of how to manipulate the transcriptional apparatus using small molecules, it would obviously be helpful to have a sophisticated understanding of how the mRNA-synthesizing machinery in eukaryotic cells works. The properties of eukaryotic transcription proteins have been reviewed exhaustively elsewhere [17,18]. Only a brief overview of the process will therefore be presented here, with a focus on understanding what activators and repressors do to modulate the efficiency of transcription.

Figure 4

A model for the transcription cycle in eukaryotic cells. When a gene is first induced, the chromatin structure must be 'loosened' (the nucleosome is shown as disappearing for simplicity; this is probably not the case), TFIID/TFIIA must associate with the core promoter and an intact holoenzyme must bind to the TFIID/TFIIA–DNA complex. This provides the first complete preinitiation complex. This species must then open the double helix to expose the template strand, then move away from the promoter and initiate mRNA synthesis. Promoter escape involves hyperphosphorylation of the carboxy-terminal domain of the polII largest subunit. TFIID/TFIIA and the mediator fragment of the holoenzyme are thought to remain at the promoter after polII and associated factors leave. To rebuild another transcription complex, only the core fragment of the holoenzyme must add. High-level transcription is the result of many reinitiation cycles. If TFIID/TFIIA or the mediator is lost before a new holoenzyme core fragment can associate, the system must fall back to some step in the initiation cycle, which is much slower. It is proposed that a major role of activators is to stabilize mediator at the promoter to facilitate multiple rounds of reinitiation.



RNA polymerase II (polII), the enzyme responsible for the transcription of all mRNA-encoding genes, is comprised of 12 polypeptides and operates in concert with a large number of general transcription factors (TFIIA, TFIIB, TFIID, etc.). Perhaps the most important of these is TFIID [19], a complex of about 13 proteins that includes the TATA-binding protein (TBP) [20,21] and TBP-associated factors (TAFs) [22]. TBP is a sequence-specific DNA-binding protein that recognizes the so-called TATA boxes (consensus: 5′-TATAAAA) present in the promoters of many genes [23,24]. One or more of the TAFs might also have DNA-binding properties [25–27]. TFIID, all of the general transcription factors and polII must assemble on the promoter to form a preinitiation complex in order to begin a transcription cycle (Figure 4). This is followed by an ATP-dependent melting of the promoter region, allowing the polymerase to associate with the template strand. Many of the protein–protein and protein–DNA interactions used to form the complex in the first place must then be severed to allow polII and

some fraction of the preinitiation complex to escape the promoter and begin their march down the gene. In this process, the elongation complex picks up a number of specialized elongation accessory factors [28] and also associates with other multiprotein complexes, for example the splicosome and the excision repair machinery. The transition from a promoter-bound to an elongating polymerase complex involves covalent modifications, in particular multiple phosphorylations of the carboxy-terminal domain (CTD) of the polII largest subunit [29]. Finally, whatever vestige of the preinitiation complex remains at the promoter must accept a new, hypophosphorylated polymerase and its attendant factors to rebuild a new preinitiation complex (Figure 4). This cycle must occur many times, because a highly active gene fires approximately every five seconds.

The complexity of the transcription machinery is daunting. The fully formed preinitiation complex has a mass greater than that of a ribosome, the network of protein–protein interactions is only partially understood, and any or all of the steps in the transcription cycle could be regulated by activators or repressors. Fortunately, recent advances in this field suggest that understanding the regulation of this process (at least at a superficial level) may not be as difficult as was feared originally. Early biochemical experiments using purified factors had suggested that formation of a preinitiation complex required a host of sequential general transcription-factor-binding events [30], leading to an almost palpable depression in the field regarding the prospects of understanding the regulation of such a complicated pathway in detail. This view has now changed dramatically with the realization that the vast majority of transcription factors travel as large, stable complexes. One does not therefore have to think about dozens of individual association steps to build a preinititation complex, as was once thought. As mentioned above, TFIID is a complex of about 13 proteins. It associates with TFIIA, comprised of three polypeptides, which helps TFIID bind to DNA, possibly by competing repressors from TBP [31,32]. Most or all of the rest of the general transcription factors, polII, and a class of proteins known as coactivators (see below) then associate in one step as parts of a huge complex known as the RNA polymerase II holoenzyme [33,34]. TFIIB, a holoenzyme component, binds to TBP [35,36], locking the components of the machine together into a single piece. It now seems likely that assembly of the preinitiation complex may require only two steps: TFIID/TFIIA–DNA binding, followed by association of the holoenzyme with this complex [37] (Figure 4).

How do activators and repressors work?

The holoenzyme is comprised of two parts. One is the so-called 'core' that includes RNA polymerase and all of the other proteins required for synthesizing mRNA. The other is the mediator [38], a complex of ~20 proteins that is required for the holoenzyme to respond to activators *in vitro* and *in vivo*. The mediator is linked to the holoenzyme through the CTD [39]. There is circumstantial evidence that this association is lost after the first firing of the promoter; polII and many associated factors move down the gene, whereas the mediator and TFIID are thought to stay behind [40,41]. This probably makes reinitiation (synthesis of transcripts 2-n) much more facile than initiation, because a stable base for formation of subsequent preinitiation complexes is present and only a fragment of the holoenzyme must reassociate. It is reasonable to assume that for highly active genes the level of mRNA synthesis is closely correlated with the number of reinitiation events for each initiation event. Once the system drops out of the reinitiation cycle as a result of loss of TFIID or mediator from the promoter, it must 'reboot' completely (Figure 4), which is probably slow. Activators clearly play an important role in reinitiation [16] and it is therefore reasonable to suggest that the major role of activators is to help to retain the mediator at the promoter during reinitiation. Some activators may also help to retain TFIID [42].

This mechanistic picture suggests that the level of transcription can be modulated by the lifetimes of the TFIID–DNA and activator–mediator complexes to give greater or lesser amounts of gene expression. The longer-lived these complexes, the more rounds of reinitiation that will occur prior to rebooting. This view is consistent with the current literature. For example, activators, such as Gal4p, that have very high affinities for mediator (C.-J. Jeong, L. Sun, S.-H. Yang, T.K. and S.A. Johnston, unpublished observations) are unusually potent activators, but only on promoters with high affinity TATA boxes. Mutations in the TATA box (Y. Xie, S.-H. Yang, L. Sun and T.K., unpublished observations) or TBP [43] that reduce the lifetime of the complex correlate directly with reduced levels of activator-mediated gene expression *in vivo*. This type of information is very important to the chemical biologist. In addition to substantiating the view of activator function presented above, the effect of point mutations provides an excellent signpost indicating which steps in a biological process should be able to be manipulated using small molecules.

Some activators may also play a role in recruiting TFIID and/or holoenzyme to promoters through direct contacts with TBP during the first initiation event [44–46]. A number of papers have also argued that activators play a major role in recruiting TFIID to promoters or maintaining it there during reinitiation through interactions with TAFs [47–49]. Several recent studies indicate that this is not a major pathway of activation *in vivo* [50–53]. Mutations that affect the rate of TFIID association with the promoter during initiation could increase the lag time between the time of induction and the onset of transcription, but would not have corresponding effects on steady-state transcription

levels. Of course, if initiation were very severely crippled, transcription would be abolished. Thus, the transcription process can be likened to a light controlled by a dimmer switch, in which the circuit must be opened for any light to be produced (initiation), but the overall output of light is controlled by a knob that can be set to any desired level (number of reinitiation events per initiation).

Of course, there are many exceptions to the above picture and no single model will describe the mechanism of action of all activators. For example, genes such as *Drosophila* Hsp70 and HIV genes activated by Tat [54,55] are clearly regulated at the level of promoter escape [56], possibly through activator-mediated recruitment of a kinase that phosphorylates the CTD and severs the association of polII with promoter-bound proteins. Small-molecule inhibitors of this kinase have been identified [57].

Activators and repressors also function at the level of chromatin structure. Chromatin is a repressed template and the first order of business in transcribing a gene must be to 'loosen up' the chromatin structure in order to promote transcription-factor binding [58–60]. The loosening of chromatin structure occurs through activator-mediated recruitment of two types of chromatin modification/remodeling complexes. One class is the histone acetyl transferases (HATs), which contain proteins that acetylate key lysine residues in the amino-terminal tails of certain histones [61]. In some way that is not yet understood, the covalent modification renders the DNA in a nucleosome far more accessible to DNA-binding transcription factors. The importance of HATs in gene activation is underscored by the recent demonstration that certain gene-specific transcriptional repressors act by recruiting a histone deacetylase complex to the target promoter, thus shutting down transcription [62,63]. This is a result that is particularly satisfying to chemical biologists, because histone deacetylase was first purified and characterized on the basis of its binding to trapoxin, a small molecule that blocks histone deacetylation *in vivo* [64]. A different type of complex that also functions at the level of chromatin is typified by the SWI/SNF chromatin remodeller in yeast [65–68] that somehow 'jiggles' core nucleosomes using energy derived from ATP hydrolysis to facilitate transcription-factor binding.

Activators can recruit these complexes in two ways. One is through direct binding. For example, the activator VP16 has been shown to bind to a protein called ADA2 [69] which is part of a multiprotein complex that also includes GCN5, a potent HAT [70]. Alternatively, there are indications that TFIID and the holoenzyme may have associated with them proteins that have HAT and/or chromatin-remodelling activity [71,72], so binding of the activation domain to one or both of these complexes may automatically recruit these activities.

As an aside, understanding how chromatin structure influences transcription is an area of tremendous opportunity for chemical biologists. No one has even the first clue regarding the structural changes in chromatin structure brought about by acetylation or ATP-dependent remodelling. Also, there are so many different HATs it seems likely that different HATs may alter chromatin structure in different ways, so the situation is almost certainly far more complicated than current models would suggest. This area of research is crying out for new probes that will allow investigators to ask and answer more detailed questions.

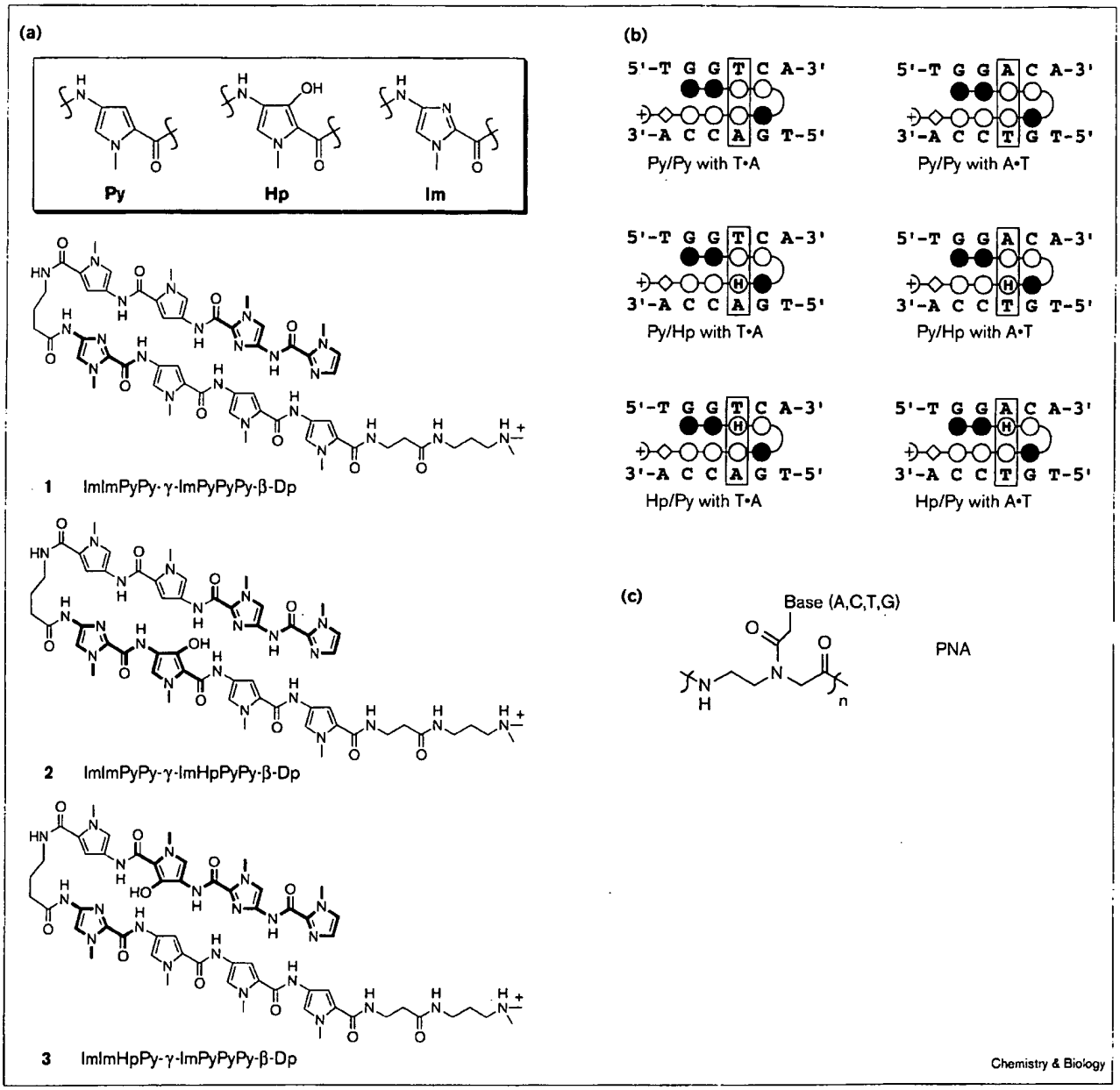
Finally, it is important to point out that the promoters of most eukaryotic genes have binding sites for more than one activator. These different proteins often interact with one another in synergistic fashion [73] and little or no gene expression results unless all of the activators are bound. Very often, this is because the proteins bind to the promoter cooperatively [74–76]. Thus, another tempting target for manipulating transcription are the activator–activator complexes that support cooperative binding.

DNA–protein interactions as molecular targets

As the examples above should have made clear, there are a number of potential strategies for manipulating the transcription process using small molecules. Of course, compounds that fundamentally alter the activity of the transcriptional machinery itself, for example an inhibitor of polII elongation, would be potent modulators of transcription but would not be gene-specific. Most strategies have therefore focused on compounds that target either the promoter of interest itself, or the activators and repressors that bind to it. The most obvious approach is to develop molecules that block activator–DNA or repressor–DNA interactions and thereby turn genes off or on artificially. Another strategy would be to find small molecules that could promote or antagonize key nuclear protein–protein interactions involved in regulation, for example between cooperating activators, between repressors and histone deacetylases, between repressors and activation domains, and possibly even between activation domains and their targets in the transcription machinery. Finally, for many genes (for example those activated by NFAT) it would be advantageous to manipulate the activity of kinases, phosphorylases or proteases that modulate the activity of an activator or repressor or control its nuclear localization. In any case, the development of protocols to design or discover small molecules that can manipulate protein–protein or protein–DNA interactions is a high priority for chemical biologists interested in manipulating gene regulation. Some particularly interesting recent advances in this area will be discussed below.

By far the most work has been carried out on compounds that bind DNA and that might serve as inhibitors of binding of proteins to overlapping sites. In particular, two types of

Figure 5



DNA-binding synthetic oligomers. **(a)** Structures of polyamides. Hp, 3-hydroxypyrrole; Im, imidazole; Py, pyrrole; β, β-alanine; γ, γ-aminobutyric acid; Dp, dimethylaminopropylamide. **(b)** Binding models for polyamides 1–3 in complex with 5′-TGGTCA-3′ and 5′-TGGACA-3′. Filled and unfilled circles represent imidazole and pyrrole rings respectively; circles containing an H represent 3-hydroxypyrrole; the curved line represents γ-aminobutyric acid; the diamond represents β-alanine; + represents the positively charged dimethyl-aminopropylamide tail group. **(c)** Basic structure of peptide nucleic acid (PNA). Parts (a,b) reproduced with permission from [86].

compounds look very promising in this role. Nielsen [77] has pioneered the development of protein nucleic acids (PNAs), which are oligomers that contain the standard purine and pyrimidine bases of an oligonucleotide but in which the sugar–phosphate backbone is replaced with a simple amide-based chain (Figure 5) [77]. PNAs bind with very high affinities to complementary single-stranded nucleic acids (both DNAs and RNAs), in fact better than standard oligonucleotides because of the lack of repulsive phosphate–phosphate interactions. Indeed, a PNA complementary to one strand of a DNA duplex will invade the double helix, pair with its complement and displace the

'like strand', forming a 'displacement loop' [78], at least in low ionic strength buffers. As might be expected, PNA invasion of a DNA duplex can abolish binding of a protein to an overlapping site and PNAs have also been employed as potent antisense agents. *In vitro*, PNAs have proven to be useful reagents for manipulating transcription and translation. Unfortunately, PNAs are not very cell-permeable, which has greatly limited their use in living cells [79]. Recently, there have been many exciting advances in moving cell-impermeable molecules through cell membranes using special peptides, however [80]. It could be that peptides conjugated to the appropriate PNA could be potent agents for manipulating gene regulation.

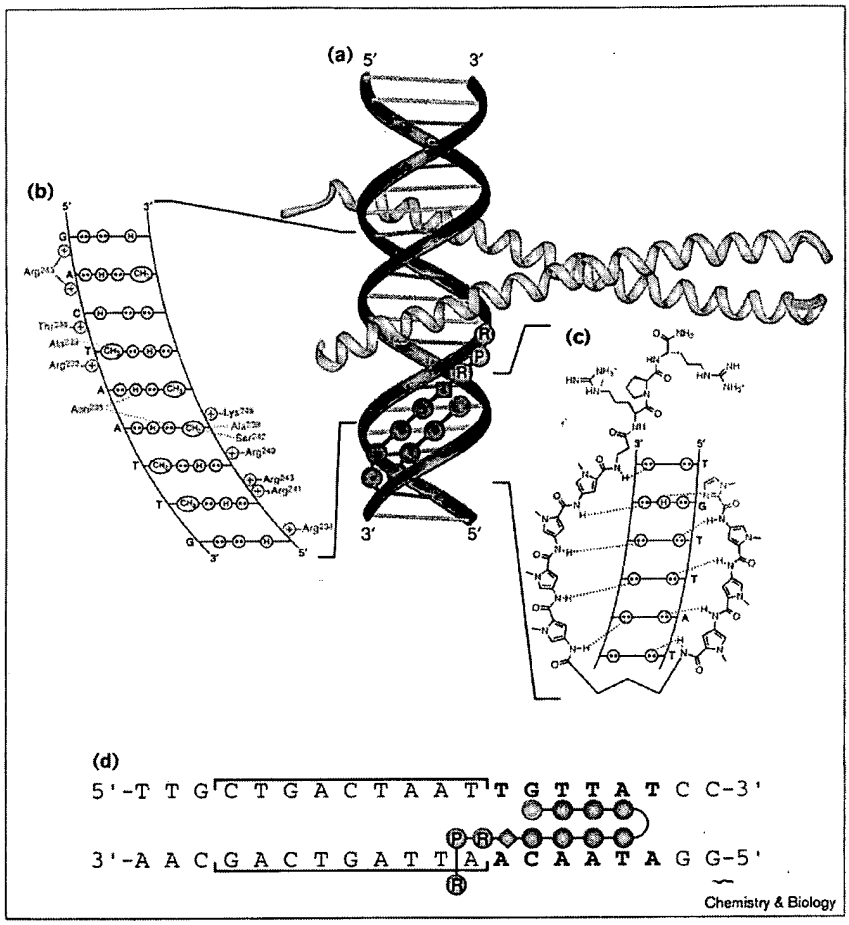
The other class of molecules that shows great promise consists of the remarkable oligomers of modified *N*-methylimidazoles and pyrroles developed by Dervan and coworkers (Figure 5) [81] (also see [82,83] for related work). These compounds were inspired by distamycin, netropsin and other minor-groove-binding natural products. It was hoped that, through both rational and some irrational experimentation, netropsin-like molecules could be made that would have greater sequence discriminatory powers than the natural products, which bind mainly A/T-rich regions. A seminal advance was the realization from nuclear magnetic resonance (NMR) experiments and other biophysical studies that two distamycins were stacked in an 'antiparallel' fashion into the minor groove in these complexes [84]. This insight led to the development of 'hairpin' oligomers that mimicked this 2:1 binding mode, with very high binding constants. Over several years, substituted imidazole and pyrrole compounds were developed that allowed recognition of any the four natural Watson–Crick base pairs in the context of this conserved structural motif [85,86]. In other words, there is now a 'code' by which a given imidazole/pyrrole pair can be selected to bind a particular base pair of DNA. An imidazole/pyrrole oligomer complementary to any sequence of double-stranded DNA can thus be designed [87,88] with little more difficulty than one would have in coming up with an oligonucleotide complementary to a piece of single-stranded DNA. This work represents a major advance in biomolecular recognition.

As one might expect, these compounds are potent inhibitors of protein–DNA interactions when minor groove contacts are critical for protein binding [89]. As most proteins make predominantly major-groove contacts, the imidazole/pyrrole oligomers will probably have to be elaborated to serve as generally useful inhibitors of sequence-specific protein–DNA interactions. There would appear to be many straightforward ways to accomplish this. For example, a recent paper [90] describes the inhibition of binding of a fragment of the yeast GCN4 protein to DNA *in vitro* using a phosphate-interference strategy. It has been known for some time that alkylation of even a single key phosphate can prevent binding of many DNA-binding proteins to their target sites, reflecting the fact that interactions between cationic or polar sidechains and the charged DNA backbone make critical contributions to the negative free energy of protein binding. With this in mind, an imidazole/pyrrole oligomer designed to bind a sequence adjacent to the recognition site of the GCN4 protein [91] was coupled to the tripeptide Arg–Pro–Arg (RPR). The hope was that the cationic arginine sidechains would be brought by the oligomer into close proximity with phosphate groups in the GCN4-recognition site, allowing the formation of strong hydrogen bonds that would occlude binding of GCN4 protein (Figure 6). In fact, this approach worked nicely *in vitro* and should be generally useful for competing the binding of many DNA-binding proteins. Other strategies might have to be explored for applications in the living cell, however, because appending charged groups to the neutral oligomers is likely to reduce their cell permeability. Indeed, the imidazole/pyrrole oligomers are sufficiently new that they have not yet really been subjected to a 'shakedown cruise' in living cells, but initial experiments look very promising [89] and they clearly hold tremendous promise as reagents for the control of gene expression. In fact, it may not be necessary to modify these compounds to manipulate the binding of proteins in the major groove of DNA in order to regulate transcription *in vivo*. This is because stable binding of TBP to TATA boxes is an important event in the transcription of a great many genes, and TBP is a minor-groove-binding protein [92]. Although it was stated above that targeting general transcription factors is a poor strategy to achieve gene-specific regulation, this is an exception because the DNA is the true target. For example, although most TATA boxes more or less resemble the consensus 5′-TATAAAA-3′, an imidazole/pyrrole oligomer could be made that recognizes only part of this site and also binds to a flanking sequence that is unique to the target promoter. As the affinity of TBP for the TATA box is very often correlated directly with transcriptional output (Y. Xie, S.-H. Yang, L. Sun and T.K., unpublished observations, also see [43,93]), manipulation of the TBP–TATA interface using the imidazole/pyrrole oligomers may allow one to modulate, rather than completely abolish, mRNA production in a highly gene-specific fashion.

There have been several other scattered reports of DNA-targeted inhibitors of specific protein binding, for example chimeras including carbohydrates and DNA-reactive small molecules [94]. Some of these appear to be quite promising and may emerge as important reagents in the future [95]. But no other class of molecules currently approaches the general utility of the PNAs and imidazole/pyrrole oligomers.

Much less work has been done on the complementary strategy for manipulating DNA–protein interactions: finding molecules that have high affinity for the DNA-binding

Figure 6



(a) A schematic model of Arg–Pro–Arg (RPR) polyamides targeted to the major groove transcription factor GCN4. **(a)** The α-helical GCN4 dimer (yellow) is shown binding to adjacent major grooves [91]. The Arg–Pro–Arg–hairpin polyamide is shown as red, blue and green balls which represent imidazole, pyrrole and Arg–Pro–Arg amino acids, respectively. The blue diamond represents β-alanine. γ-Aminobutyric acid is designated as a curved line. **(b)** The contacts between one GCN4 monomer and the major groove of one half-site of 5′-CTGACTAAT-3′ are depicted (adapted from [91]). Circles with two dots represent the lone pairs of the N7 of purines, the O4 of thymine and the O6 of guanine. Circles containing an H represent the N6 and N4 hydrogens of the exocyclic amines of adenine and cytosine, respectively. The C5 methyl group of thymine is depicted as a circle with CH_3 inside. Protein sidechains that make hydrogen bonds or van der Waals contacts to the bases are shown in purple and connected to the DNA via a dotted line. Green and purple plus signs represent protein residues that electrostatically contact the phosphate backbone. The residues that are predicted to be disrupted by an Arg–Pro–Arg polyamide are shown in green. **(c)** The hydrogen-bonding model of the eight-ring hairpin polyamide ImPyPyPy-γ-PyPyPyPy-β-RPR bound to the minor groove of 5′-TGTTAT-3′. Circles with two dots represent the lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogens of guanines. Putative hydrogen bonds are illustrated by dotted lines. Py and Im rings are represented as blue and red rings, respectively. The Arg–Pro–Arg moiety is green. **(d)** The model of the polyamide binding its target site (bold) adjacent to the GCN4 binding site (brackets). Polyamide residues are as in (a). Reproduced from [90].

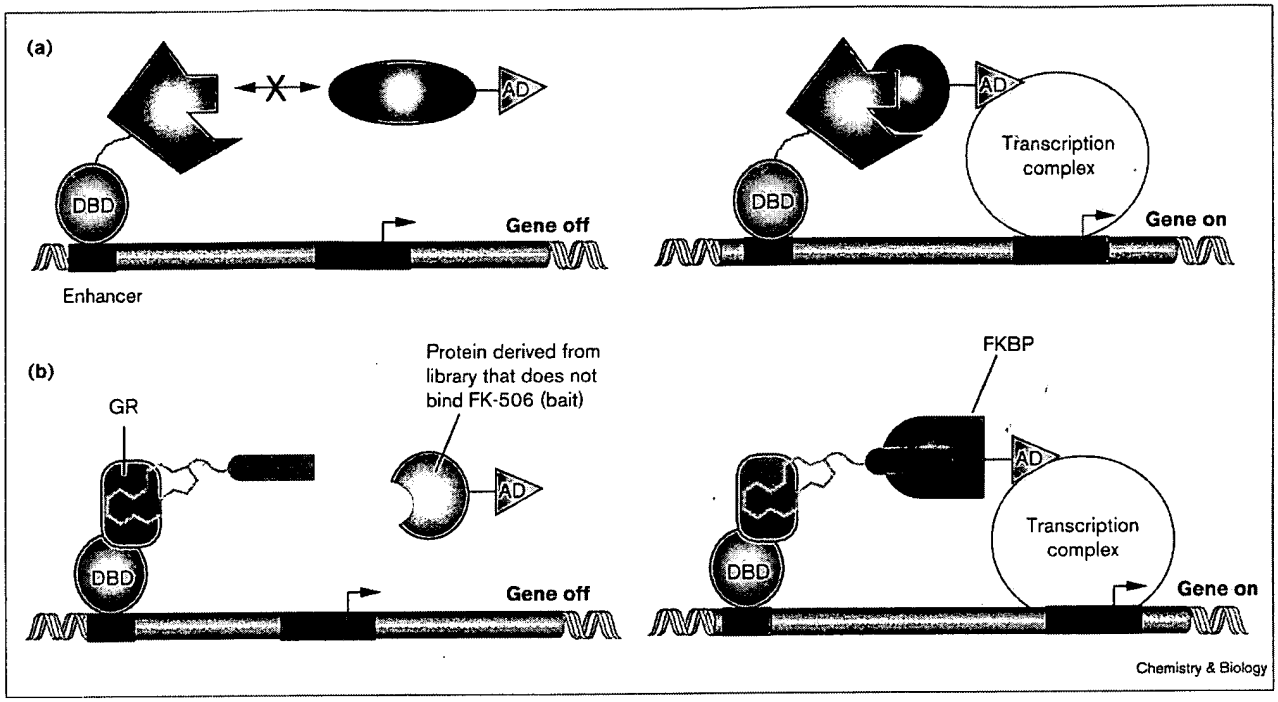
domains of key activators or repressors and therefore block their association with DNA. This is generally considered to be an even harder task than DNA recognition. Although broad structural families of DNA-binding domains certainly exist, polypeptide targets lack a single, well-defined architecture, which is a hallmark of the DNA double helix. Nonetheless, we predict that this approach will be a growth area in the future as chemical biologists begin to learn how to make molecules that bind specific peptide and protein sequences.

Targeting protein–protein interactions: better ways to find a needle in a haystack

Finding new molecular matchmakers or disrupters of protein–protein interactions is a very high priority for chemical biologists. Unnatural molecules that have these properties have been very hard to come by. Part of the reason for this is that pharmaceutical companies, where most protein-binding synthetic molecules come from, have traditionally concentrated on developing enzyme inhibitors rather than manipulators of protein–protein interactions. This will almost certainly change. Once these efforts are brought up to speed it will be critical to already have general assays by which libraries, combinatorial or otherwise, can be screened for molecules that have the property of interest, because rational design is unlikely to succeed in most cases. Although it is increasingly common to screen libraries for molecules that bind a given target protein, finding a matchmaker or disrupter is a difficult process because only a fraction of the molecules that bind a particular protein will influence its interaction with other factors.

Recently, there has been important progress in the design of high-throughput screens or selections designed to identify

Figure 7

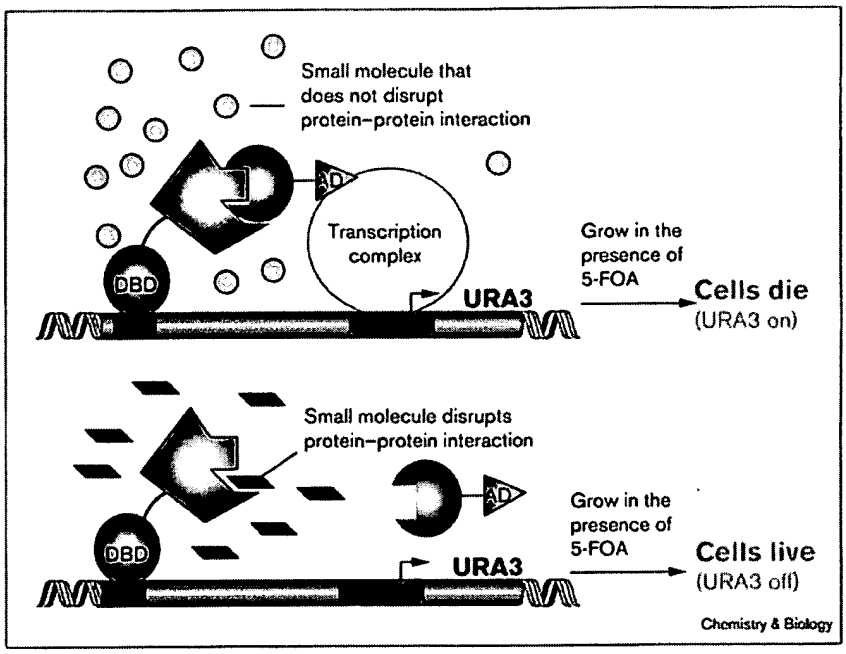


(a). Schematic representation of the two-hybrid system, a genetic method used to detect protein–protein interactions. If the proteins fused to a DNA-binding domain and an activation domain do not interact, then transcription of a reporter gene will be very low. If these proteins do interact, however, then a functional activator will be reconstituted and the reporter gene will be expressed at high levels. **(b).** The three-hybrid assay to detect proteins that bind a given small molecule. The orange steroid-shaped symbol represents dexamethasone. The red bullet represents FK-506. GR, glucocorticoid receptor. See text for details.

those molecules with the desired matchmaker or disrupter function. Much of this work was inspired by systems set up by geneticists (who have been practising combinatorial chemistry of a sort for much longer than chemists have) to identify proteins that interact with one another. Generically known as two-hybrid assays [96], this family of methods takes advantage of the fact that, in many promoter contexts, the DNA-binding and activation domains of an activator function more or less independently of one another (see [97,98] for exceptions), but must be physically connected. For example, if the Gal4-activation and DNA-binding domains are severed and these fragments are expressed in a yeast strain deleted for wild-type GAL4, no transcription of Gal4p-targeted genes will occur. If the genes encoding proteins X and Y are fused to the DNA encoding the severed GAL4 domains, and X and Y bind to one another, Gal4p activity will be reconstituted and transcription of the target genes will occur (Figure 7a). To make this system more convenient, strains have been constructed in which activated transcription of a target gene is essential for cell survival, making the process a straightforward selection for protein–protein interactions. Using this approach, it is now routine to screen genomic cDNA libraries fused to the activation domain for genes or gene fragments that encode polypeptides which bind to a particular 'bait' protein fused to the DNA-binding domain [99].

Many variations of this basic strategy have been reported for more specialized applications. Most relevant to this discussion are the 'three-hybrid' system and the 'reverse two-hybrid' system. The first, reported by Licitra and Liu [100], is a clever method to identify the protein targets of biologically active natural products. The technique employs the same strategy of reconsituting the activity of a severed transcriptional activator, but is designed such that a small molecule must bridge the interaction between the proteins fused to the DNA-binding and activation domains (Figure 7b). In a proof of principle experiment, the rat glucocorticoid receptor (GR) hormone-binding domain was fused to a sequence-specific DNA-binding domain and a cDNA library was fused to an activation domain. The screen was then carried out in the presence of a chimeric small molecule consisting of dexamethasone (a GR ligand) linked to FK-506. As expected, a screen for cells in which a reporter gene was activated resulted in the isolation of the gene encoding FKBP. This demonstrates the feasibility of using genetic screens for probing small-molecule–protein interactions *in vivo*. The reverse

Figure 8



The reverse three-hybrid system for detecting small molecules that disrupt a protein–protein interaction. See text for details.

two-hybrid system is a method to select for mutations that abrogate protein–protein interactions [101,102]. In this case, the 'reporter' gene targeted by the reconstituted activator is chosen such that its expression is toxic and therefore can be selected against.

Schreiber and coworkers [103] have recently combined elements of the two-hybrid, three-hybrid and reverse two-hybrid systems to create a convenient system for identifying small molecule disruptors of protein–protein interactions. Their approach is shown in Figure 8. As in the reverse two-hybrid system, yeast cells were engineered so that a protein–protein interaction which reconstitutes activator function is conditionally toxic. This was accomplished using a standard yeast genetics trick of placing the URA3 gene under the control of the severed activator and growing the cells in the presence of 5-fluoroorotic acid (5-FOA). When operated on by the URA3 gene product, 5-FOA is transformed into a toxic substance but in the absence of URA3 expression it is harmless. Alternatively, expression of URA3 is nontoxic in the absence of 5-FOA, allowing clones that contain interacting proteins to be grown and propagated easily. The fusion proteins containing the DNA-binding domain and activation domain were placed under the control of the Gal4 protein. Gal4p-mediated expression is essentially zero when the cells are grown in glucose, but occurs at high levels in galactose-containing media. Thus, both the expression of the interacting proteins and the consequences of their interaction can be controlled by the experimenter. The utility of this system was demonstrated by taking advantage of the fact that FK-506 inhibits the binding of FKBP to the transforming growth factor β type I receptor R1 [103]. As expected, growth of yeast containing R1 fused to a DNA-binding domain and FKBP fused to an activation domain was sensitive to the presence of 5-FOA, but this sensitivity was abrogated by FK-506.

Our laboratory has developed a different genetic assay (based on a method originally devised by Hu and coworkers [104]) in which two different fusion proteins, each containing the λ repressor DNA-binding domain, are expressed in *Escherichia coli* equipped with a repressor-controlled green fluorescent protein (GFP) gene. The fusion proteins lack the normal dimerization domain of the Repressor. If the proteins fused to the DNA-binding domain interact and artificially dimerize the repressor DNA-binding domain, GFP expression is therefore blocked (C. Ackerson and T.K., unpublished observations). If the fusion proteins do not heterodimerize, or if a molecule is present that blocks the interaction of the proteins fused to the repressor fragment, however, then GFP is expressed at high levels. These bright green cells are easy to identify in a background of dark cells.

'Spray and pray' and the 'squeege': combining the power of combinatorial libraries and genetic assays

The biological screens and selections described above are ideal for high-throughput screening protocols in which cells are introduced into the wells of 96 well (or denser)

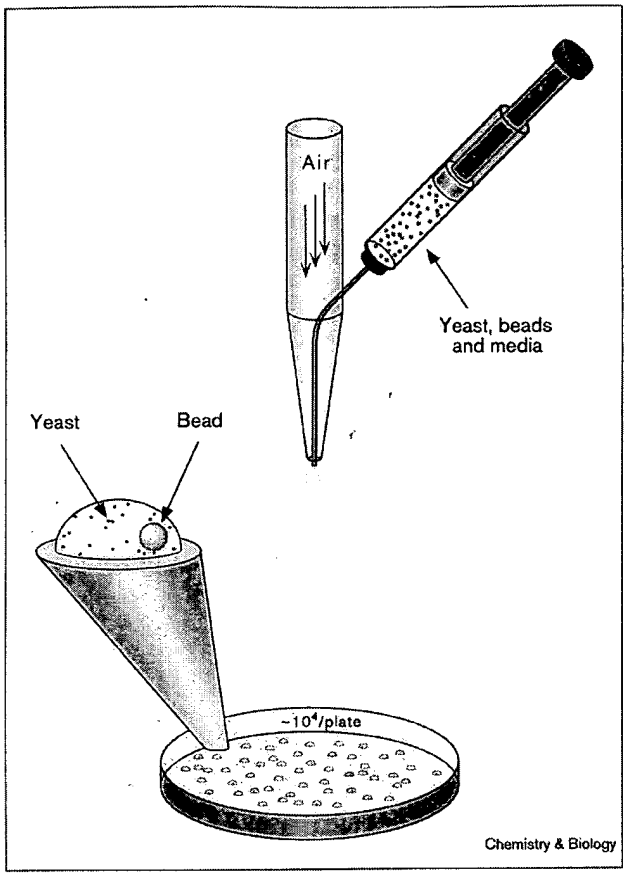
microtiter plates, each of which contains a different chemical. In this way, the entire suite of compounds possessed by a pharmaceutical company could be screened for matchmaker or disrupter activity in a reasonable period of time. Of potentially even greater use, however, would be to apply these techniques to screening combinatorial libraries made by the split and pool method [105] in which each bead is derivatized with many copies of a unique compound. The trick here would be to somehow expose the *E. coli* or yeast reporter strain to many, many different beads in a spatially segregated manner so that ideally one yeast cell sees one bead in some kind of microincubator where the chemical compound can be released from the bead. This knotty problem has been solved elegantly in two ways. One, called the 'stochastic nanodroplet' method [106], employs the simple idea of mixing yeast cells and chemically modified beads together then spraying them as a fine mist onto an agar plate (Figure 9). If the flow and levels of yeast and beads are controlled appropriately, the 'nanodroplets' sprayed onto the plate will contain from zero to a few beads (or 0–1 if bead density is kept very low) per droplet as well as one to a few yeast cells. The nanodroplets are now spatially segregated on the plate. If yeast growth is unimpeded, each nanodroplet will give rise to a yeast colony. Borchardt *et al.* [106] used beads linked to rapamycin via a photocleavable linker to demonstrate that when the plates were photolyzed enough toxic rapamycin was released from the bead to diffuse into the yeast cells and strongly inhibit growth. In theory, the same approach could be employed using combinatorial libraries of compounds attached to the beads by the same photolabile linker and a yeast or *E. coli* reporter strain engineered to report on the state of a particular protein–protein interaction.

The second approach [107] also employs spatially segregated nanodroplets as mini-incubators. In this case, however, the bead/cell mixture is layered onto a plastic plate with small wells that are extremely closely packed. These plates are produced by a photolithographic/imprinting technique and precoated with a substance that makes the bottom of the wells cell-adherent. Again, the amounts of beads and cells are chosen so that after the excess liquid is 'squeegeed' off the plate, the nanodroplets that remain have one to a few beads and a few cells in them. The advantage of this technique is that the squeegee procedure is much more gentle than the spraying technique and even much more fragile mammalian cells can be used in this format. This combination of genetic selection and combinatorial chemistry technologies promises to be an extremely effective route to the discovery of small molecule disrupters and matchmakers.

Synthetic mimics of activators and repressors

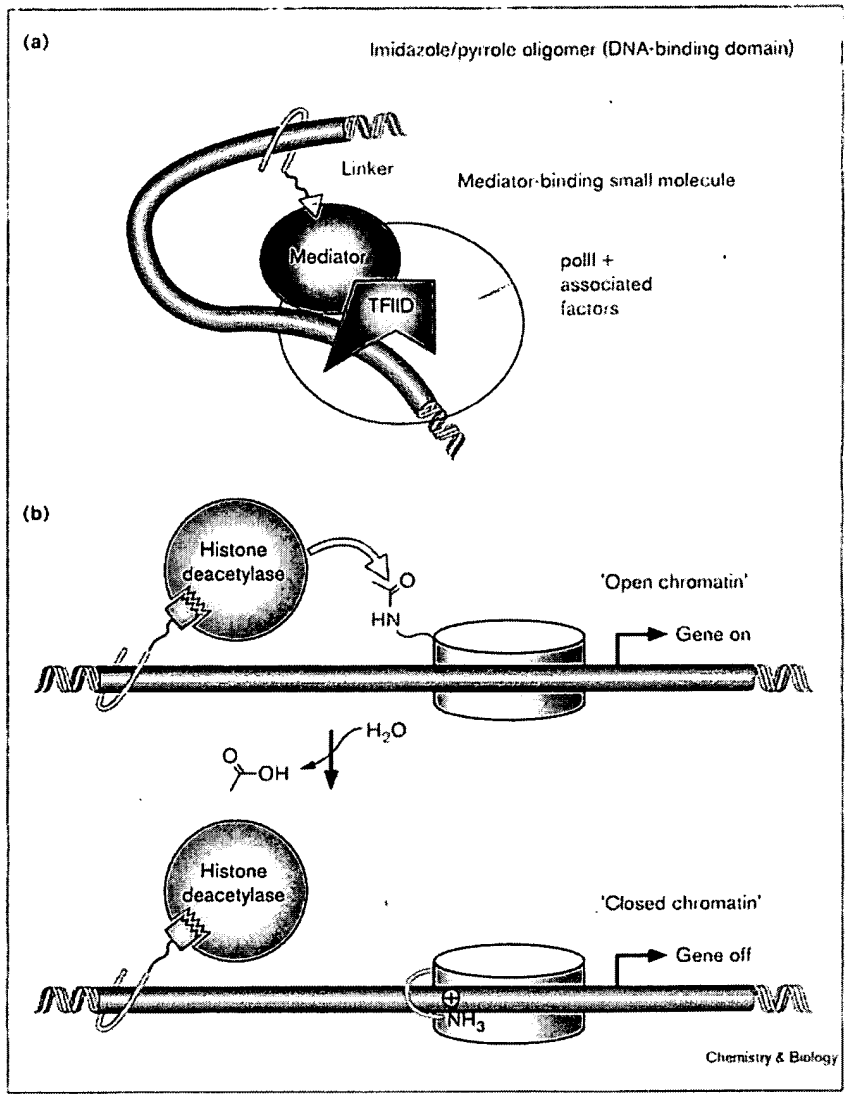
Although almost all of the above discussion has focused on using small molecules to manipulate the interactions of transcription factors with each other and with DNA, perhaps the ultimate goal in this area is to make cell-permeable small molecules that directly mimic the activity of repressors or activators. Such molecules would be extremely valuable research tools and potentially revolutionary drugs. For example, a large percentage of human cancers are associated with a defective p53 gene that encodes a transcriptional activator important in regulating cell-cycle progression [108]. If a nontoxic small molecule could be made that would activate the transcription of the p53 target genes, the impact on human health would be enormous. Although this idea might have seemed to be pure fantasy a decade

Figure 9



Formation of nanodroplets by spraying. A mixture of beads evenly dispersed in medium containing yeast is slowly injected into a stream of air forming a fine mist. When layered on to a surface such as a Petri dish this forms into nanodroplets. The average volume of the droplets is controlled by the amount of liquid applied to the surface. For a droplet volume of 50–200 nl it is possible to deposit 5000–8000 droplets in the area of a Petri dish (80 cm^2). The fraction of droplets containing beads depends on the density of beads in the medium prior to spraying. When a mixture of 80 μm Tentagel beads and medium are sprayed at a density of 14,000 beads/ml, approximately 10% of the droplets contain beads. This results in 1000 bead-containing droplets per Petri dish. Of the bead-containing droplets we find that 88% contain a single bead, 10% contain two beads, 1.3% contain three beads, and 0.7% contain four beads. Reproduced from [106].

Figure 10



Proposed scheme to make completely synthetic activators and repressors. **(a)** A synthetic activator could be constructed from a Dervan-type pyrrole/imidazole oligomer targeted to a sequence just upstream of the target gene. Fused to the artificial DNA-binding domain would be a small molecule selected to bind tightly to the mediator fragment of the polII holoenzyme. This should act as an artificial activation domain. **(b)** A synthetic repressor could be constructed by fusing a pyrrole/imidazole oligomer to a small molecule that binds to, but does not inhibit, histone deacetylase. This would result in a highly inaccessible template in the region around the small-molecule-binding site, thereby strongly repressing transcription.

ago, some of the advances in our understanding of transcriptional regulatory mechanisms suggest that the development of such a mimic is now eminently feasible. Transcriptional regulators appear mainly to be matchmakers between specific DNA sequences (promoters) and either the transcriptional machinery itself or catalytic activities that condense or decondense the chromatin structure. Making synthetic mimics of transcriptional regulatory proteins should therefore be orders of magnitude simpler than making small molecules with catalytic activities comparable to enzymes (for an intriguing study directed towards the creation of an artificial coactivator, see [109]).

As described above, the major role of many activators is probably to recruit the polII holoenzyme to the promoter and, perhaps more importantly, retain the mediator fragment there through many rounds of transcription. In theory, one could therefore make a synthetic activator by linking a sequence-specific DNA-binding molecule, for example the appropriate imidazole/pyrrole oligomer, with a molecule selected to bind to a surface-exposed mediator constituent (Figure 10). At least in yeast, most of the mediator components have been identified and the genes cloned [110], so this is quite feasible. Comparable information on the human mediator will undoubtedly be available in the near future.

Similarly, the recent discovery that many repressors function mainly to recruit a histone deacetylase complex to a given gene suggests a straightforward method to make an

artificial repressor. Again an imidazole/pyrrole oligomer could be used to localize a covalently linked histone-deacetylase-binding molecule isolated from a library. In this case, the mammalian histone deacetylase is known [64] (see discussion above), and it would not surprise us if exactly this sort of experiment is underway is several laboratories. In fact, we predict that, if the imidazole/pyrrole oligonucleotides, or possibly PNAs, prove to be generally useful *in vivo* (i.e., artificial DNA-binding domains are readily available), within 5–10 years biological chemists will have in hand an arsenal of small, cell-permeable molecules with which they can artificially control the expression of a very significant fraction of genes in the human genome. These are exciting times.

Acknowledgements

We thank Stephen Johnston (UT-Southwestern) and members of the Kodadek laboratory for many helpful conversations. Work in this laboratory on *E. coli*-based screening technology was supported by a grant from the Welch Foundation and studies of yeast transcription were supported by grants from the American Cancer Society and the NIH.

References

1. Brent, R. & Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA binding specificity of a prokaryotic repressor. *Cell* **43**, 729-736.
2. Kodadek, T. & Johnston, S.A. (1995). The dangers of 'splicing and dicing': on the use of chimeric transcriptional activators *in vitro*. *Chem. Biol.* **2**, 187-194.
3. Lohr, D., Venkov, P. & Zlatanova, J. (1995). Transcriptional regulation in the yeast GAL gene family: a complex genetic network. *FASEB J.* **9**, 777-787.
4. Johnston, S.A., Salmeron, J.M. Jr. & Dincher, S.S. (1987). Interaction of positive and negative regulatory proteins in the galactose regulation of yeast. *Cell* **50**, 143-146.
5. Leuther, K.K. & Johnston, S.A. (1992). Nondissociation of GAL4 and GAL80 *in vivo* after galactose induction. *Science* **256**, 1333-1335.
6. May, M.J. & Ghosh, S. (1997). Rel/NF-kappa B and I kappa B proteins, an overview. *Semin. Cancer Biol.* **8**, 63-73.
7. Ho, S., *et al.* & Crabtree, G.R. (1996). The mechanism of action of cyclosporin A and FK-506. *Clin. Immunol. Immunopathol.* **80**, S40-S45.
8. Schreiber, S.L. & Crabtree, G.R. (1992). The mechanism of action of cyclosporin A and FK506. *Immunol. Today* **13**, 136-42.
9. Van, D.G., Standaert, R.F., Karplus, P.A., Schreiber, S.L. & Clardy, J. (1991). Atomic structure of FKBP-FK506, an immunophilin-immunosuppressant complex. *Science* **252**, 839-42.
10. Ke, H., *et al.* & Walsh, C.T. (1994). Crystal structures of cyclophilin A complexed with cyclosporin A and *N*-methyl-4-[(E)-2-butenyl]-4,4-dimethylthreonine cyclosporin A. *Structure* **2**, 33-44.
11. Liu, J., *et al.* & Schreiber, S.L. (1991). Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807-815.
12. Shaw, K.T.Y., *et al.* & Hogan, P.G. (1995). Immunosuppressive drugs prevent a rapid dephosphorylation of transcription factor NFAT1 in stimulated immune cells. *Proc. Natl Acad. Sci. USA* **92**, 11205-11209.
13. Loh, C., *et al.* & Rao, A. (1996). Calcineurin binds the transcription factor NFAT1 and reversibly regulates its activity. *J. Biol. Chem.* **271**, 10884-10891.
14. Crabtree, G.R. & Schreiber, S.L. (1996). Three-part inventions: intracellular signalling and induced proximity. *Trends Biochem. Sci.* **21**, 418-422.
15. Spencer, D.M., Wandless, T.J., Schreiber, S.L. & Crabtree, G.R. (1993). Controlling signal transduction with synthetic ligands. *Science* **262**, 1019-1024.
16. Ho, S.N., Biggar, S.R., Spencer, D.M., Schreiber, S.L. & Crabtree, G.R. (1996). Dimeric ligands define a role for transcriptional activation domains in reinitiation. *Nature* **382**, 822-826.
17. Orphanides, G., Lagrange, T. & Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev.* **10**, 2657-2683.
18. Conaway, R.C. & Conaway, J.W. (1993). General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.* **62**, 161-190.
19. Burley, S.K. & Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Ann. Rev. Biochem.* **65**, 769-799.
20. Kim, J.L., Nikolov, B.D. & Burley, S.K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**, 520-527.
21. Kim, Y., Geiger, J.H., Hahn, S. & Sigler, P.B. (1993). Crystal structure of a yeast TBP/TATA complex. *Nature* **365**, 512-520.
22. Dynlacht, B.D., Hoey, T. & Tjian, R. (1991). Isolation of coactivators associated with the TATA binding protein that mediate transcriptional activation. *Cell* **66**, 563-576.
23. Hoopes, B.C., LeBlanc, J.F. & Hawley, D. (1992). Kinetic analysis of yeast TFIID–TATA box complex formation suggests a multi-step pathway. *J. Biol. Chem.* **267**, 11539-11547.
24. Coleman, R.A., & Pugh, B.F. (1995). A kinetic mechanism by which the TATA-binding protein achieves sequence specific DNA binding. *J. Biol. Chem.* **270**, 13850-13859.
25. Kaufmann, J. & Smale, S. (1994). Direct recognition of initiator elements by a component of the transcription factor IID complex. *Genes Dev.* **8**, 821-829.
26. Verrijzer, C.P., Yokomori, K., Chen, J.-L. & Tjian, R. (1994). *Drosophila* TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* **264**, 933-941.
27. Verrijzer, C.P., Chen, J.-L. Yokomori, K. & Tjian, R. (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* **81**, 1115-1125.
28. Reines, D., Conaway, J.W. & Conaway, R.C. (1996). The RNA polymerase II general elongation factors. *Trends Biochem. Sci.* 351-355.
29. O'Brien, T., Hardin, S., Greenleaf, A. & Lis, J.T. (1994). Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* **370**, 75-77.
30. Burtowski, S., Hahn, S., Guarente, L. & Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**, 549-561.
31. Geiger, J.H., Hahn, S., Lee, S. & Sigler, P.B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272**, 830-836.
32. Kokubo, T., Swanson, M.J., Nishikawa, J.-I., Hinnebusch, A.G. & Nakatani, Y. (1998). The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Mol. Cell. Biol.* **18**, 1003-1012.
33. Thompson, C.M. Koleske, A.J. Chao, D.M. & Young, R.A. (1993). A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**, 1361-1375.
34. Koleske, A.J. & Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**, 466-469.
35. Nikolov, D.B., *et al.* & Burley, S.K. (1995). Crystal structure of a TFIIB–TBP–TATA element ternary complex. *Nature* **377**, 119-128.
36. Burley, S.K. (1996). Picking up the TAB. *Nature* **381**, 112-113.
37. Stargell, L.A. & Struhl, K. (1996). Mechanisms of transcriptional activation *in vivo*: two steps forward. *Trends Genet.* **12**, 311-315.
38. Bjorkland, S. & Kim, Y.-J. (1996). Mediator of transcriptional regulation. *Trends Biochem. Sci.* **21**, 335-337.
39. Kim, Y.-L., Bjorklund, S., Li, Y., Sayre, M.H. & Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599-608.
40. Svejstrup, J.Q., *et al.* & Kornberg, R.D. (1997). Evidence for a mediator cycle at the initiation of transcription. *Proc. Natl Acad. Sci. USA* **94**, 6075-6078.
41. Zawel, L., Kumar, K.P. & Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* **9**, 1479-1490.
42. Lieberman, P.M. & Berk, A.J. (1991). The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction. *Genes Dev.* **5**, 2441-2454.
43. Lee, M. & Struhl, K. (1995). Mutations on the DNA-binding surface of TATA-binding protein can specifically impair the response to acidic activators *in vivo*. *Mol. Cell. Biol.* **15**, 5461-5469.
44. Klein, C. & Struhl, K. (1994). Increased recruitment of TATA-binding protein to the promoter by transcriptional activation domains *in vivo*. *Science* **266**, 280-282.
45. Barberis, A., *et al.* & Ptashne, M. (1995). Contact with a component of the polymerase II holoenzyme suffices for gene activation. *Cell* **81**, 359-368.
46. Farrell, S., Simkovich, N., Yu, Y., Barberis, A. & Ptashne, M. (1996). Gene activation by recruitment of the RNA polymerase II holoenzyme. *Genes Dev.* **10**, 2358-2367.
47. Chen, J.-L., Attardi, L.D., Verrijzer, C.P., Yokomori, K. & Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* **79**, 93-105.

48. Sauer, F., Hansen, S.K. & Tjian, R. (1995). Multiple TAF_{II}s directing synergistic activation of transcription. *Science* **270**, 1783-1788.
49. Verrijzer, C.P. & Tjian, R. (1996). TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci.* **9**, 338-342.
50. Moqtaderi, Z., Bai, Y., Poon, D., Weil, P.A. & Struhl, K. (1996). TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature* **383**, 188-191.
51. Apone, L.M., Virbasius, C.A., Reese, J.C. & Green, M.R. (1996). Yeast TAF_{II}90 is required for cell-cycle progression through G2/M but not for general transcription activation. *Genes Dev.* **10**, 2368-2380.
52. Walker, S.S., Reese, J.C., Apone, L.M. & Green, M.R. (1996). Transcription activation in cells lacking TAF_{II}s. *Nature* **383**, 185-188.
53. Walker, S.S. Shen, W.-C. Reese, J.C. Apone, L.M. & Green, M.R. (1997). Yeast TAF_{II}145 required for transcription of G1/S cyclin genes and regulated by the cellular growth state. *Cell* **90**, 607-614.
54. Cujec, T.P., *et al.* & Peterlin, B.M. (1997). The HIV transactivator Tat binds to the CDK-activating kinase and activates the phosphorylation of the carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 2645-2657.
55. Garcia-Martinez, L.F., *et al.* & Gaynor, R.B. (1997). Purification of a Tat-associated kinase reveals a $TF_{II}H$ complex that modulates HIV-1 transcription. *EMBO J.* **16**, 2836-2850.
56. Gilmour, D.S. & Lis, J.T. (1986). RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster*. *Mol. Cell. Biol.* **6**, 3984-3989.
57. Mancebo, H.S.Y., *et al.* & Flores, O. (1997). P-TEF kinase is required for HIV Tat transcriptional activation *in vivo* and *in vitro*. *Genes Dev.* **11**, 2633-2644.
58. Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. *Nature* **355**, 219-224.
59. Kingston, R.E., Bunker, C.A. & Imbalzano, A.N. (1996). Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev.* **10**, 905-920.
60. Kadonaga, J.T. (1998). Eukaryotic transcription: An interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92**, 307-313.
61. Brownell, J.E. & Allis, C.D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr. Opin. Genet Dev.* **6**, 176-184.
62. Brehm, A., *et al.* & Kouzardis, T. (1998). Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597-601.
63. Magnaghi-Jaulin, L., *et al.* & Harel-Bellan, 'A. (1998). Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391**, 601-605.
64. Taunton, J., Hassig, C.A. & Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408-411.
65. Burns, L.G. & Peterson, C.L. (1997). The yeast SWI–SNF complex facilitates binding of a transcriptional activator to nucleosomal sites *in vivo*. *Mol. Cell. Biol.* **17**, 4811-4819.
66. Quinn, J., Fyrberg, A.M., Ganster, R.W., Schmidt, M.C. & Peterson, C.L. (1996). DNA-binding properties of the SWI/SNF complex. *Nature* **379**, 844-847.
67. Cote, J., Quinn, J., Workman, J.L. & Peterson, C.L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **265**, 53-60.
68. Kwon, H., Imbalzano, A.N., Khaviri, P.A., Kingston, R.E. & Green, M.R. (1994). Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. *Nature* **370**, 477-481.
69. Berger, S.L., *et al.* & Guarente, L. (1992). Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic actviation domains. *Cell* **70**, 251-265.
70. Candau, R., Zhou, J.X., Allis, C.D. & Berger, S.L. (1997). Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function *in vivo*. *EMBO J.* **16**, 555-565.
71. Mizzen, C.A., *et al.*, & Allis, C.D. (1996). The TAFII250 subunit of TFIID has histone acetyltransferase activity. *Cell* **87**, 1261-1270.
72. Wilson, C.J., *et al.* & Roung, R.A. (1996). RNA polymerase II holoenzyme contains Swi/Snf regulator involved in chromatin remodeling. *Cell* **84**, 235-244.
73. Herschlag, D. & Johnson, F.B. (1992). Synergism in transcriptional activation: a kinetic view. *Genes Dev.* **7**, 173-179.
74. Kim, T.K. & Maniatis, T. (1997). The mechanism of transcriptional synergy of an *in vitro* assembled interferon-β enhanceosome. *Mol. Cell* **1**, 119-129.
75. Giese, K., Kingsley, C., Kirshner, J.R. & Grosschedl, R. (1995). Assembly and function of a TCRα enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* **9**, 995-1008.
76. Molkentin, J.D., *et al.* & Olson, E.N. (1998). A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* **93**, 215-228.
77. Nielsen, P.E. (1997). Design of sequence-specific DNA-binding ligands. *Chem. Eur. J.* **3**, 505-508.
78. Footer, M., Egholm, M., Kron, S., Coull, J.M. & Matsudaira, P. (1996). Biochemical evidence that a D-loop is part of a four-stranded PNA–DNA bundle. Nickel-mediated cleavage of duplex DNA by a Gly–Gly–His–Bis–PNA. *Biochemistry* **35**, 10673-10679.
79. Good, L. & Nielsen, P.E. (1998). Antisense inhibition of gene expression in bacteria by PNA targeted to mRNA. *Nat. Biotechnol.* **16**, 355-358.
80. Rojas, M., Donahue, J.P., Tan, Z. & Lin, Y.-Z. (1998). Genetic engineering of proteins with cell membrane permeability. *Nat. Biotechnol.* **16**, 370-375.
81. Trauger, J.W., Baird, E.E. & Dervan, P.B. (1996). Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **382**, 559-561.
82. Bruice, T.C., Mei, H.-Y., He, G.-X. & Lopez, V. (1992). Rational design of tripyrrole peptides that complex DNA by both selective minor groove binding and electrostatic interaction with the phosphate. *Proc. Natl Acad. Sci. USA* **89**, 1700-1704.
83. Chiang, S.-Y., Bruice, T.C., Azizkhan, J.C., Gawron, L. & Beerman, T.A. (1997). Targeting E2F1-DNA complexes with microgonotropen DNA binding agents. *Proc. Natl Sci. USA* **94**, 2811-2816.
84. Mrksich, M., *et al.* & Wemmer, D.E. (1992). Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc. Natl. Acad. Sci. USA* **89**, 7586-7590.
85. Geierstanger, B.H., Mrksich, M., Dervan, P.B. & Wemmer, D.E. (1994). Design of a G.C-specific DNA minor groove-binding peptide. *Science* **266**, 646-650.
86. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. & Dervan, P.B. (1998). Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468-471.
87. Wemmer, D.E. & Dervan, P.B. (1997). Targeting the minor groove of DNA. *Curr. Opin. Struct. Biol.* **7**, 355-361.
88. White, S., Baird, E.E. & Dervan, P.B. (1997). On the pairing rules in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem. Biol.* **4**, 569-578.
89. Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E. & Dervan, P.B. (1997). Regulation of gene expression by small molecules. *Nature* **387**, 202-205.
90. Bremer, R.E., Baird, E.E. & Dervan, P.B. (1998). Inhibition of major groove-binding proteins by pyrrole-imidazole polyamides with an Arg–Pro–Arg patch. *Chem. Biol.* **5**, 119-133.
91. Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted helices. *Cell* **71**, 1223-1237.
92. Starr, D.B. & Hawley, D.K. (1991). TFIID binds in the minor groove of the TATA box. *Cell* **67**, 1231-1240.
93. Harbury, P.A.B. & Struhl, K. (1989). Functional distinctions between yeast TATA elements. *Mol. Cell. Biol.* **9**, 5298-5304.
94. Ho, S.N., Boyer, S.H., Schreiber, S.L., Danishefsky, S.J. & Crabtree, G.R. (1994). Specific inhibition of formation of transcription complexes by a calicheamicin oligosaccharide: a paradigm for the development of transcriptional antagonists. *Proc. Natl Acad. Sci. USA* **91**, 9203-9207.
95. Kahne, D. (1995). Strategies for the design of minor groove binders: a re-evaluation based on the emergence of site-selective carbohydrate binders. *Chem. Biol.* **2**, 7-12.
96. Fields, S. & Song, O.-k. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.
97. Vashee, S. & Kodadek, T. (1995). The activation domain of GAL4 protein mediates cooperative promoter binding with general transcription factors *in vivo*. *Proc. Natl Acad. Sci. USA* **92**, 10683-10687.
98. Vashee, S., Melcher, K., Ding, W.V., Johnston, S.A. & Kodadek, T. (1998). Evidence for two modes of cooperative DNA binding *in vivo* that do not involve direct protein-protein interactions. *Curr. Biol.* **8**, 452-458.
99. Phizicky, E.M. & Fields, S. (1995). Protein–protein interactions: Methods for detection and analysis. *Microb. Rev.* **59**, 94-123.
100. Lecitra, E.J. & Liu, J.O. (1996). A three-hybrid system for detecting small ligand-protein interactions. *Proc. Natl Acad. Sci. USA* **93**, 12817-12821.

101. Vidal, M., Braun, P., Chen, E., Boeke, J.D. & Harlow, E. (1996). Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. *Proc. Natl Acad. Sci. USA* **93**, 10321-10326.
102. Leanna, C.A. & Hannink, M. (1996). The reverse two-hybrid system: a genetic scheme for selection against specific protein/protein interactions. *Nucleic Acids Res.* **24**, 3341-3347.
103. Huang, J. & Schreiber, S.L. (1997). A yeast genetic system for selecting small molecule inhibitors of protein-protein interactions in nanodroplets. *Proc. Natl Acad. Sci. USA* **94**, 13396-13401.
104. Hu, J., O'Shea, E.K., Kim, P.S. & Sauer, R.T. (1990). Sequence requirements for coiled-coils: Analysis with λ repressor-GCN4 leucine zipper fusions. *Science* **250**, 1400-1403.
105. Borman, S. (1997). Combinatorial chemistry. *Chem. Eng. News* **75**, 43-62.
106. Borchardt, A., Liberles, S.D., Biggar, S.R., Crabtree, G.R. & Schreiber, S.L. (1997). Small molecule-dependent genetic selection in stochastic nanodroplets as a means of detecting protein–ligand interactions on a large scale. *Chem. Biol.* **4**, 961-968.
107. You, A.J., Jackman, R.J., Whitesides, G.M. & Schreiber, S.L. (1997). A miniaturized arrayed assay format for detecting small molecule–protein interactions in cells. *Chem. Biol.* **4**, 969-975.
108. Cox, L.S. & Lane, D.P. (1995). Tumour suppressors, kinases and clamps: how p53 regulates the cell cycle in response to DNA damage. *Bioessays* **17**, 501-508.
109. Nyanguile, O. Uesugi, M. Austin, DJ. & Verdine, G.L. (1997). A nonnatural transcriptional coactivator. *Proc. Natl Acad. Sci. USA* **94**, 13402-13406.
110. Myers, L.C., *et al.* & Kornberg, R.D. (1998). The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **12**, 45-54.

Cellular Microbiology (2002) 4(8), 471–482

Technoreview

Using small molecules to study big questions in cellular microbiology

Gary E. Ward,[1]* Kimberly L. Carey[1] and Nicholas J. Westwood[2†]
[1]*Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, Vermont 05405, USA.*
[2]*Institute of Chemistry and Cell Biology and The Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA.*

Summary

High-throughput screening of small molecules is used extensively in pharmaceutical settings for the purpose of drug discovery. In the case of antimicrobials, this involves the identification of small molecules that are significantly more toxic to the microbe than to the host. Only a small percentage of the small molecules identified in these screens have been studied in sufficient detail to explain the molecular basis of their antimicrobial effect. Rarer still are small molecule screens undertaken with the explicit goal of learning more about the biology of a particular microbe or the mechanism of its interaction with its host. Recent technological advances in small molecule synthesis and high-throughput screening have made such mechanism-directed small molecule approaches a powerful and accessible experimental option. In this article, we provide an overview of the methods and technical requirements and we discuss the potential of small molecule approaches to address important and often otherwise experimentally intractable problems in cellular microbiology.

Small molecules and small molecule approaches

Small organic molecules that target specific proteins and thereby act as either agonists or antagonists of particular cellular processes have played a major role in cell biological research. Consider how much has been learned about the inner workings of eukaryotic cells through the use of small molecules such as cytochalasin D, brefeldin A or tunicamycin, which disrupt actin polymerization, intracellular trafficking, and N-glycosylation respectively (Cooper, 1987; McDowell and Schwarz, 1988; Ohmori and Toyama, 1992; Jackson, 2000 – see Fig. 1 for structures). Because the specific targets of many of these small molecules have been well established, the observation that a particular small molecule induces or inhibits a particular biological process can often be used to implicate the known target of that molecule in the process. For example, it is now well established that actin plays a central role in the intra- and intercellular motility of *Listeria monocytogenes* (reviewed in Cameron *et al.*, 2000), in the formation of 'pedestals' on intestinal epithelial cells by enteropathogenic *E. coli* (reviewed in Celli *et al.*, 2000) and in the invasion of host cells by *Toxoplasma gondii* (Dobrowolski and Sibley, 1996). One of the earliest observations implicating actin in each of these processes was the inhibitory effect of cytochalasin D (Ryning and Remington, 1978; Schwartzman and Pfefferkorn, 1983; da Silva *et al.*, 1989; Tilney and Portnoy, 1989). Small molecules with well-characterized targets have likewise been used to study signalling events during host–pathogen interaction (e.g. Rodriguez *et al.*, 1995; Kenny and Finlay, 1997; Pelkmans *et al.*, 2002), parasite cytoskeletal function (Morrissette and Sibley, 2002) and intracellular trafficking pathways in infected cells (e.g. Hackstadt *et al.*, 1996; Coppens *et al.*, 2000).

Where do small molecules such as these come from? Typically, they have been identified in, and isolated from, complex mixtures of natural products by following a biological activity (Clark, 1996; Ohizumi, 1997; Grabley and Thiericke, 1999). An alternative approach, which is rapidly gaining momentum, is to search large, structurally diverse (Martin, 2001) collections of *individual* small molecules for those that cause a desired biological effect. The availability of small molecule collections in the required format (see below) has eliminated the time- and labour-intensive steps needed to purify bioactive small molecules from complex mixtures. While it remains a challenge to match the molecular diversity present in nature through synthetic approaches, reproducibility and access to methods for

Received 24 January, 2002; revised 1 May, 2002; accepted 2 May, 2002. *For correspondence: E-mail gward@zoo.uvm.edu; Tel. (+1) 802 656 4868; Fax (+1) 802 656 8749. †Present address: Department of Chemistry and the Centre for Biomolecular Sciences, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST, Scotland, UK.

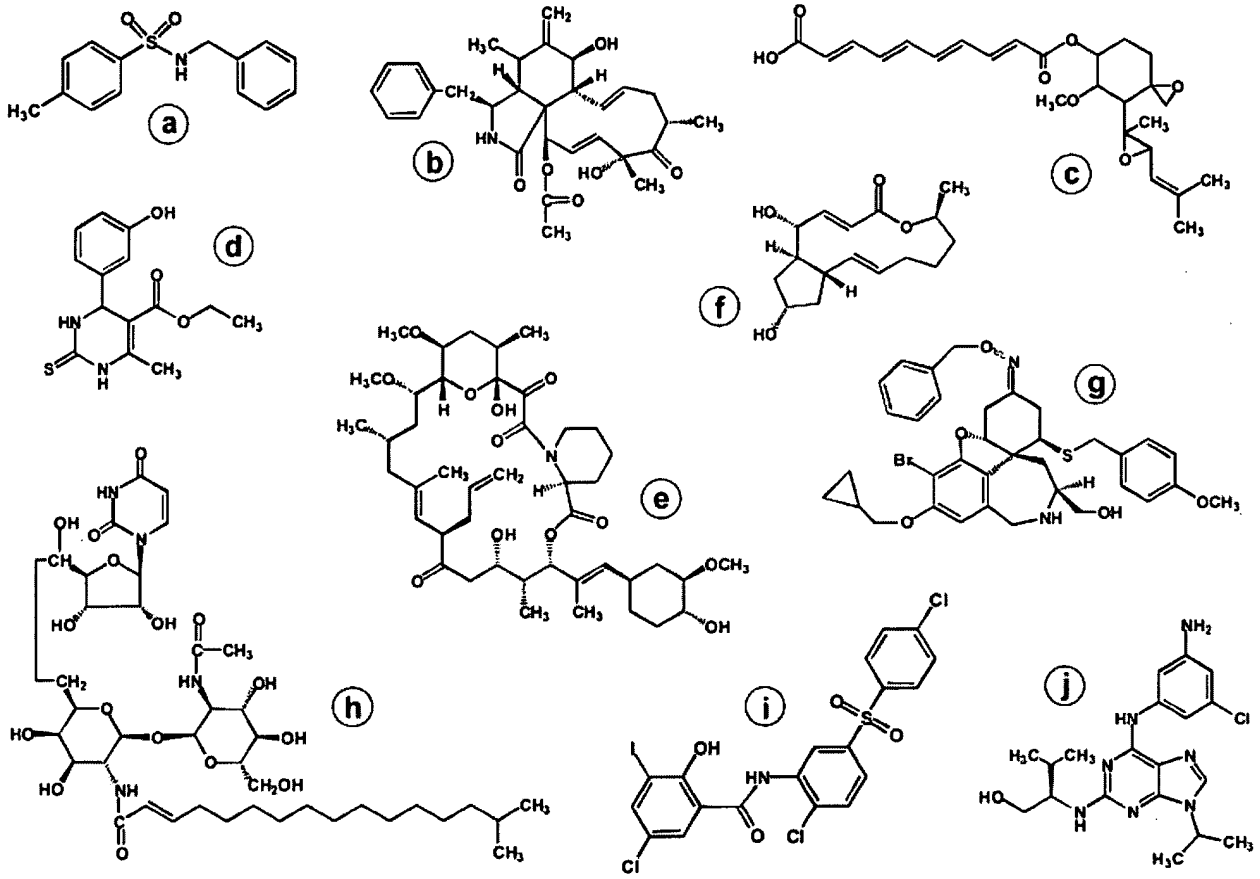


Fig. 1. What is a 'small molecule?' The term small molecule will mean different things to different people. In the context of the target- and phenotype-based screens discussed here, small molecules are organic, non-peptide compounds (for a recent review of peptide and peptidomimetic drugs, see Al-Obeidi *et al.*, 1998), typically <1500 Da. They are either synthetic or derived from natural product extracts. A key structural feature is often a rigid multiring core structure that reduces the entropic cost paid on binding of the small molecule to a protein. Membrane permeability is frequently, though not always (e.g. Quillan *et al.*, 1995), an important property. Examples of bioactive small molecules referred to in the text include: (a) BTS – inhibits skeletal muscle myosin II by weakening its interaction with actin (Cheung *et al.*, 2002); (b) cytochalasinD – inhibits actin polymerization through direct interaction with actin (Cooper, 1987; Ohmori and Toyama, 1992); (c) fumagillin – inhibits angiogenesis and binds to the methionine aminopeptidase MetAP-2 (Sin *et al.*, 1997); (d) monastrol – arrests cells in mitosis with monoastral spindles, through inhibition of the mitotic kinesin Eg5 (Mayer *et al.*, 1999; Kapoor *et al.*, 2000); (e) FK506 – inhibits the protein phosphatase calcineurin through direct interaction with a 12 kDa FK506-binding protein (Schreiber and Crabtree, 1992); (f) brefeldin A – inhibits protein trafficking by binding to and stabilizing a transient complex between ADP-ribosylation factor-1 and its guanine nucleotide exchange factor (Peyroche *et al.*, 1999; Chardin and McCormick, 1999); (g) secramine – blocks trafficking of proteins from the Golgi to the plasma membrane, target unknown (Pelish *et al.*, 2001); (h) tunicamycin – inhibits protein glycosylation via a direct effect on *N*-acetylglucosamine transferases (McDowell and Schwarz, 1988); (i) BH3I-2′ – induces apoptosis by binding to the BH-3-binding pocket of Bcl-x_L (Degterev *et al.*, 2001) (j) aminopurvalanol – induces leukaemic cell differentiation, binds to cyclin-dependent kinase 1 (Rosania *et al.*, 1999).

rapid structural optimization account for the increasing interest in this approach.

Collections of individual small molecules can be used in two distinctly different ways to study cell biological mechanisms. In the 'phenotype'-based approach (Fig. 2), a collection of small molecules is screened for those that affect a particular biological process. The bioactive molecules (or derivatives of these molecules) are then used as reagents to identify the cellular components that function in that process. This approach is analogous to a classical forward genetic screen in which a collection of random mutants is screened for those that exhibit a particular phenotype. Alternatively, a collection of small molecules can be screened for those that alter the activity of a single purified or recombinant protein; the identified small molecules are then added to intact cells to ascertain their biological effects. This 'target'-based approach (Fig. 2) is analogous to a reverse genetic strategy. Because of the conceptual similarities to forward and reverse genetic strategies, these small molecule approaches have been referred to as 'chemical genetics' (Mitchison, 1994; Schreiber, 1998). The small molecule approaches differ from rational drug design (Setti and Micetich, 1996; Klebe, 2000) in that they require no detailed structural informa-

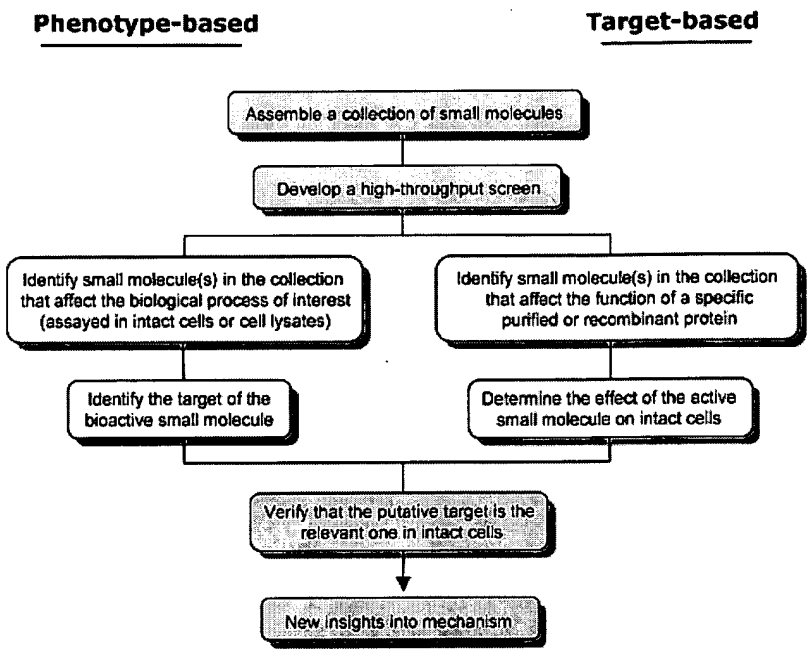


Fig. 2. Summary of the phenotype- and target-based small molecule approaches discussed in this review. Phenotype-based approaches can be used to identify proteins involved in a specific cellular process, while target-based approaches are used to elucidate the cellular functions of known proteins. In both cases, novel 'protein-small molecule' pairs of physiological relevance are identified, and new small molecules that perturb a particular cellular process are developed.

tion about the target molecule ahead of time. Rather, through sampling large numbers of structurally diverse small molecules, an appropriately designed screen allows the protein target to 'select' active structures (Crews and Splittgerber, 1999).

Small molecule approaches are essentially pharmacological in origin. The overriding issue in all pharmacological studies is one of *specificity*, and many small molecule agonists/antagonists have proven to be highly specific. For example, cytochalasin D inhibits the motility of both mammalian cells and parasitic protozoa through a direct effect on actin, and resistance to cytochalasin D is conferred in both cases by amino acid substitutions that map to similar positions on the actin monomer (Ohmori and Toyama, 1992; Dobrowolski and Sibley, 1996). A recently described small molecule inhibitor of skeletal muscle myosin II [BTS: see Fig. 1, structure (a)] shows little inhibition of myosins from other tissues and cell types, and the inhibitory activity of BTS is remarkably sensitive to small changes in its structure (Cheung *et al.*, 2002).

Recent phenotype-based screens have led to the identification of a number of small molecules with significant potential as cell biological probes. For example, in a screen of 16 320 small molecules for agents that disrupt mitotic spindle formation, a small molecule was identified that specifically inhibited the mitotic kinesin Eg5 (Mayer *et al.*, 1999). This small molecule, named 'monastrol' [see Fig. 1, structure (d)] because it leads to the formation of mitotic spindles with a single aster, has since been used to provide novel insights into the mechanism of mitotic spindle assembly (Kapoor *et al.*, 2000). Phenotype-based approaches have also been used to successfully identify new small molecule agonists and/or antagonists of cell cycle progression (reviewed in Stockwell, 2000), apoptosis (Degterev *et al.*, 2001), leukaemic cell differentiation (Rosania *et al.*, 1999), zebrafish development (Peterson *et al.*, 2000), Sir2-mediated transcriptional silencing in yeast (Grozinger *et al.*, 2001), membrane trafficking and secretion (Yamaguchi *et al.*, 1999; Feng *et al.*, 2001; Pelish *et al.*, 2001), centrosome duplication (Mayer *et al.*, 2001), actin assembly (Peterson *et al.*, 2001), and cell surface receptor activation (Tian *et al.*, 1998; Zhang *et al.*, 1999). Recent successes of target-based screening include the identification of specific inhibitors/activators of a variety of kinases, phosphatases, tumor suppressors, myosins, ion channels, and cell surface and steroid receptors (reviewed in Stockwell, 2000; see also Gray *et al.*, 1998; Foster *et al.*, 1999; Cheung *et al.*, 2001; 2002; Mattheakis and Savchenko, 2001; Shen *et al.*, 2001).

Why use small molecules rather than classical genetics?

Genetic approaches have proven to be an extremely powerful way to dissect the mechanisms underlying complex cellular processes. However, there are situations in which a small molecule approach may be more useful than standard forward or reverse genetics. First, for many biologically interesting organisms, including a disproportionate number of pathogens (in particular obligate intracellular pathogens), standard genetic tools are either unavailable or rudimentary. Studying the function of essential or recessive genes in such organisms can be problematic. Second, even in cases where dominant

negative alleles, RNA interference, or conditional (e.g. temperature sensitive) mutations are available, genetic approaches are generally not well-suited to studying dynamic cell biological processes that occur on a time scale of seconds to minutes. The addition and removal of small molecules allows protein function to be perturbed at specific times and in a controlled manner, which will usually be a more informative way to study such phenomena. Finally, and of particular importance in the field of cellular microbiology, using small molecules to investigate basic biological mechanisms has the direct advantage that it may simultaneously yield promising leads for the development of new antimicrobial drugs. This is particularly relevant in the case of third world pathogens, which have been all but ignored by the pharmaceutical industry due to a lack of financial incentive (Werbovetz, 2000).

In this review, we will focus on phenotype-based small molecule approaches, their technical requirements and their potential to address important mechanistic questions in cellular microbiology. We will pay particular attention to one of the more challenging aspects of the approach: identifying the targets of small molecules determined to be bioactive.

What is required?

Establishing and screening a small molecule collection using high-throughput screening (HTS) methods was, until recently, only feasible in pharmaceutical companies. However, as access to collections and HTS technology has improved (Selzer *et al.*, 2000), the academic research community has become increasingly interested in these approaches (Gura, 2000). The cost of assembling the requisite technology and compound collections is now within the reach of many academic institutions and/or collaborative units within those institutions. Funding agencies such as the National Cancer Institute (NCI) have formally recognized the value and potential of small molecule approaches through a variety of programs and initiatives, including the Discovery Services of the Developmental Therapeutics Program (Weinstein *et al.*, 1997), Molecular Targets Drug Discovery grants, and the Molecular Target Laboratories initiative (see *Science* 2002, 295: 1991).

Small molecule collections

The first step in any small molecule screen (Fig. 2) is to assemble a collection of highly pure, chemically diverse (Martin, 2001) small molecules on a scale (number and quantity) that is compatible with the available screening technology. Small molecule collections are available to qualified investigators from the National Cancer Institute (http://dtp.nci.nih.gov/webdata.html) and the National Institute of Neurological Disorders and Stroke [the NINDS Custom Collection, distributed by Microsource Discovery Systems (Gaylordsville, CT)]. A number of companies also provide large (up to 250 000 member) collections of small molecules, including 'tailored' subsets of the main collection if required (see: *http://www.combichem.net/suppliers/compound.html*). These collections are supplied as individual dry films/powders or as stock solutions in DMSO either in 96- or 384-well microtitre plates. With the increased availability of 'off the shelf' collections, the most challenging decision has become which collection or subset of small molecules to investigate. Researchers focused on the discovery of drugs are guided by the increasingly defined concept of a 'drug-like molecule' (Muegge *et al.*, 2001). Commercial collections may be skewed towards such compounds, having weeded out molecules with undesirable pharmacokinetic properties. Researchers looking to increase understanding of biological mechanisms using the small molecule approach have less strictly defined criteria, and are arguably only limited by their funding, screening capacity and personal bias.

An increasingly important criterion in selecting a collection is access to additional structural analogs of the bioactive members of the collection. In the phenotype-based approach discussed here, structural optimization or derivitization of the bioactive small molecule is frequently an essential component of target identification strategies. A molecular tool kit that contains structurally related small molecules varying in affinity for the same protein target aids both affinity chromatography and *in vivo* target confirmation studies (see below). For non-chemists using these approaches, an important challenge is to identify suitable collaborators with synthetic expertise or commercial suppliers with an ability to carry out affordable synthetic follow-up work. These approaches are inherently multidisciplinary, and broadening the training base of the scientists engaged in the work will ultimately be the most productive way forward.

A powerful method of addressing at least some of the synthetic challenges posed is to take advantage of the under-exploited technique of 'split-pool' organic synthesis (see Fig. 3). This technique allows researchers to construct 'libraries' of small molecules that contain from hundreds to millions of *structurally related* small molecules (Dolle, 2001). Provided the core structure used in a particular library is relevant to the biological question being addressed, rapid access to a large number of analogs is assured. With much of the technology for split-pool synthesis now firmly established (e.g. Tallarico *et al.*, 2001; Walling *et al.*, 2002), the forefront of research in this area has returned to the inherent chemical challenges.

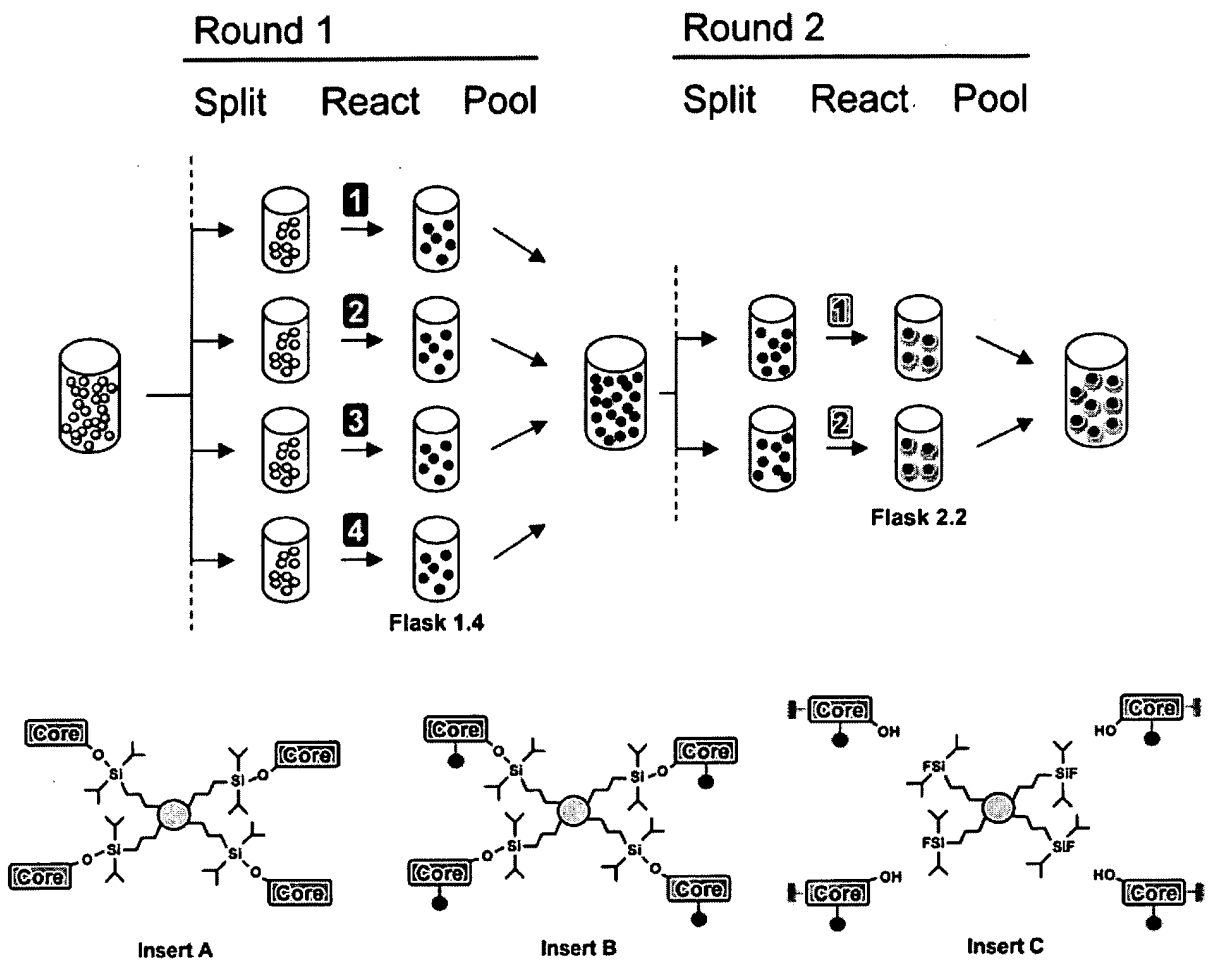


Fig. 3. Split-Pool Synthesis. One method of preparing large numbers of structurally related small molecules is to use split-pool methods (Furka *et al.*, 1991; Tan and Burbaum, 2000). The synthesis starts with a set of beads each having multiple copies of the same core small molecule attached to it (grey beads in the top diagram; see also Insert A). The beads are then split out into any number of different reaction flasks (four in this example). Analogous chemical reactions are used to attach a different building block to the small molecules in each flask (e.g. a purple building block is added to all copies of the small molecule on each bead in flask 1.4; Insert B). The beads from each flask are then pooled together and split back out into a second set of reaction flasks. Only two flasks are shown but any number may be used. As in Round 1, analogous chemistry is used in each reaction flask to add a second building block to all small molecules on each bead. The beads from each flask are then pooled again, and two rounds of split-pool synthesis have been completed.The key strength of split-pool synthesis is that small molecules containing all possible combinations of building blocks are prepared in just a few chemical steps, each combination being attached to a different bead. In the example shown, eight possible combinations are prepared in six chemical steps. If 10 reaction flasks/building blocks had been used in Rounds 1 and 2 and two further rounds of split-pool synthesis were carried out, each with 10 building blocks, all possible combinations ($10 \times 10 \times 10 \times 10 = 10\,000$ small molecules) of the 40 (10×4) building blocks could be prepared in just 40 chemical steps. Beads may be retained at each stage of the process to facilitate rapid resynthesis of bioactive molecules. Further details and a discussion of encoding strategies can be found elsewhere (Ohlemeyer *et al.*, 1993; Tan and Burbaum, 2000; Affleck, 2001 and Kassel, 2001). Insert A: Schematic representation of a solid phase synthesis bead. A linker system covalently attaches the core small molecule to the bead. Many copies of the core molecule are attached to each bead (e.g. approximately 6×10^{16} copies per 550 µm polystyrene bead; Tallarico *et al.*, 2001); only four are depicted here. Insert B: Schematic representation of beads from flask 1.4, each of which has multiple copies of the modified small molecule attached to it. Modification occurs by covalent attachment of a purple building block to the same functional group in each starting small molecule. Insert C: At the end of the synthesis the beads are physically separated from each other (e.g. Walling *et al.*, 2002) and independently treated with a reagent that cleaves the small molecules from the bead for analytical and biological testing. The bead shown was present in flasks 1.4 (purple building block) and 2.2 (light blue building block).

High-throughput screening

Once a collection of small molecules has been assembled, a high-throughput screening method capable of identifying molecules with the desired activity within the collection is required. The methods used to screen vary widely, depending on the question under investigation (e.g. see Lam *et al.*, 1997; Rose *et al.*, 1996), but a few generalizations can be made (Rademann and Jung, 2000; Selzer *et al.*, 2000; Stockwell, 2000). The screen must be robust, as automated as possible, and miniaturized to the extent possible in order to keep costs down

and maintain precious small molecule stocks (Kariv and Chung, 2001). Both 384- and 1536-well microtitre plates are commercially available and are typically the format of choice (Kariv and Chung, 2001; Walling *et al.*, 2002). Despite their small volume (20–70 μl per well for 384-well plates, 2–10 μl for 1536-well plates), they support the growth of mammalian cells (Stockwell *et al.*, 1999). Pin array devices can be used to transfer nanolitre to microlitre volumes of liquid between multiwell plates, either manually or using a robotic system (Walling *et al.*, 2002). A variety of liquid handling devices for transferring small volumes of solutions (5–50 μl) into and out of multiwell plates are also commercially available (Brush, 1999; Selzer *et al.*, 2000; Kariv and Chung, 2001). Negative and (when available) positive controls should be included on each plate, and all 'hits' from the primary screen should be independently reproduced with aliquots individually picked from the library plates or, preferably, with reordered (in the case of commercial collections) or resynthesized material. Confirmation of the chemical structure of the selected small molecules and active or inactive analogs is necessary due to the possibility of degradation on storage (Cheung *et al.*, 2001).

Many target-based screens and some phenotype-based assays can be readily achieved in plate-reader format, using absorbance, luminescence, radioactivity, scintillation proximity or various fluorescence-based assays as a readout (Kenny *et al.*, 1998; Cortese, 2000; Selzer *et al.*, 2000; Blake, 2001; Kariv and Chung, 2001). Alternatively, if a specific antibody is available for measuring the protein or process of interest, a versatile high-throughput assay called the 'cytoblot' is a powerful screening approach (Mayer *et al.*, 1999; Stockwell *et al.*, 1999).

Not all cell biological processes of interest lend themselves to immunological, biochemical or other standard readouts. This is particularly true for phenotype-based screens, which may involve the analysis of complex phenomena such as embryonic development (Peterson *et al.*, 2000), changes in cell morphology (Rosania *et al.*, 2000; Yarrow *et al.*, 2000) or changes in intracellular trafficking patterns (Feng *et al.*, 2001; Pelish *et al.*, 2001). In cases like these, the inspection of images collected by light (transmitted or fluorescence) microscopy is typically the only readout available and is often rich in information (Blake, 2001). Special microtitre plates with thin polystyrene or glass bottoms are commercially available and suitable for imaging (Ward and Carey, 1999). Plate handling and image collection can be automated, either by custom modification of existing microscope equipment or with automated microscopes specifically designed for this purpose [available from such suppliers as Universal Imaging (Downington, PA), Cellomics (Pittsburg, PA), or Applied Imaging (Santa Clara, CA)].

Automated image acquisition generates enormous amounts of data. For example, in a recent dual immunfluorescence-based screen for small molecule inhibitors of *Toxoplasma* invasion (N. J. Westwood *et al.* in preparation), we typically processed 18 (384-well) plates per day, generating over 9Gb of digital image data. The challenge in an automated, microscope-based screen therefore becomes how to archive, retrieve and analyse the data collected (Kenny *et al.*, 1998). Data archival and retrieval can be handled using commercially available software packages specifically designed for this purpose [e.g. ActivityBase (ID Business Solutions; Cambridge, MA); RS^3 Discovery (Accelrys/Pharmacopeia; Princeton, NJ); MDL Screen (MDL Information Systems; San Leandro, CA)]. Automated analysis of parameters such as object size, number or brightness is relatively straightforward (e.g. see Mayer *et al.*, 2001), and more sophisticated analysis tools are available within most image analysis software packages [available from such suppliers as Universal Imaging (Downington, PA); Media Cybernetics (Silver Spring, MD); and Improvision (Lexington, MA)].

It should be noted that while automated image acquisition and analysis is more time-efficient and can be less prone to investigator error and/or bias than manual screening, the hardware and software to fully automate image acquisition and analysis are expensive. Manual, microscope-based screens enjoy a rich and productive history within the field of genetics (e.g. Hartwell *et al.*, 1974; Nusslein-Volhard and Wieschaus, 1980; Driever *et al.*, 1996; Haffter *et al.*, 1996), and manual small molecule screening should likewise be considered when the phenotype under investigation does not lend itself well to automated analysis (e.g. Peterson *et al.*, 2000; Rosania *et al.*, 2000), when cost is an issue, or when screens are undertaken on a relatively small scale (e.g. Rosania *et al.*, 1999). Manual screening may be particularly convenient for the secondary screening of compounds identified as active in a primary screen (e.g. Mayer *et al.*, 1999) or for screening second generation compound collections designed around hit compounds (e.g. Brady *et al.*, 1998).

Approaches to target identification

The guess-and-test approach

The rate-limiting step in most phenotype-based approaches is the identification of the molecular targets of compounds determined to be active in the high-throughput screen. The simplest approach to target identification is to guess what the target might be and test the hypothesis. This approach should not be overlooked amongst the more elegant and systematic strategies available, as it can lead to target identification more rapidly than any of the other approaches. Educated

guesses can be based on: (a) the observed biological effect of the small molecule; (b) the known effects (or lack thereof) of the small molecule in other screens; and (c) a structure-based literature search to determine whether the small molecule or structurally related derivatives have pharmacological activity in any other system. The identification of Eg5 as the molecular target for monastrol (see above) is a particularly good example of the use of this approach (Mayer *et al.*, 1999). Although the NCI's recently announced 'Chembank' initiative to develop a database of small molecules and their effects on gene products, pathways and phenotypes (Adam, 2001) will eventually simplify searching for known pharmacological activities of individual small molecules, significant in-house databases already exist for certain commercial libraries. Broader, structure-based searches of the chemical literature are also possible, through such portals as SciFinder Scholar (Chemical Abstracts Service, American Chemical Society). The guess-and-test approach will become increasingly powerful as databases documenting the phenotypic consequences of systematic gene disruption in model organisms are further expanded (Delneri *et al.*, 2001; Kim, 2001).

Biochemical and cDNA expression-based approaches

Several biochemical methods have been used for target identification. The most common method has been to synthesize a derivative of the active molecule containing both a detectable 'tag' (radioactive or non-radioactive, such as biotin) and, if necessary, a photoactivatable cross-linking group. The modified small molecule is then covalently cross-linked to the target in cells or cell extracts, and the tag is used to follow the target during standard biochemical purification (Sin *et al.*, 1997; Meng *et al.*, 1999). Alternatively, the small molecule can be coupled to a solid phase matrix via an appropriately designed 'handle' (Mitchison, 1994) and used to affinity purify the target from extracts (e.g. Rosania *et al.*, 1999; 2000; Knockaert *et al.*, 2000; 2002). In 'drug-westerns', tagged small molecule derivatives are used to probe either electrophoretically resolved cell extracts or cDNA expression libraries (Tanaka *et al.*, 1999). Other potential cDNA expression-based approaches include phage display and the 'three hybrid' technique (see King, 1999; Stockwell, 2000 and references therein). Transcriptional profiling is a useful method for identifying the targets of small molecules that affect gene expression (e.g. Marton *et al.*, 1998; Hughes *et al.*, 2000).

A recently developed cDNA expression-based approach with great potential is the protein microarray, in which collections of individual recombinant proteins are spotted onto glass slides at high spatial density (MacBeath and Schreiber, 2000; Zhu *et al.*, 2001). These microarrays can be probed with tagged small molecules to identify their binding partners. Snyder and colleagues recently expressed 93.5% of all the possible open reading frames in the yeast genome, purified the individual fusion proteins, and arrayed them all onto a single glass microscope slide (Zhu *et al.*, 2001). This technical tour-de-force clearly demonstrates the feasibility of genome-wide searches for small molecule-binding proteins.

While each of the above approaches has been successfully used to identify small molecule targets, each also has its limitations. Biochemical approaches are problematic if the target molecules are either present in low abundance or difficult to extract. cDNA expression-based approaches will fail if the target is not present in the cDNA library or if the expressed target protein is improperly folded. cDNA expression-based approaches may also fail (as will drug-westerns) if the target protein requires interaction with other proteins or lipids for small molecule binding.

Genetic approaches

A powerful approach to target identification which does not require structural modification of the bioactive small molecule is to generate (e.g. by chemical mutagenesis) mutant cells/organisms resistant to the small molecule and then identify specific gene products in these mutants that are able to confer resistance when transfected back into wild-type (sensitive) cells. Sequencing the corresponding mutant and wild-type alleles identifies both a candidate target protein and the mutation that underlies resistance. Further biochemical experiments are required to prove that the identified gene product is the direct target for the small molecule, rather than an indirect effector. The approach will not work if it proves impossible to generate resistance, if the tools to do complementation cloning are unavailable, or if the mutant gene product is unable to confer resistance in the presence of the wild-type gene product. Nonetheless, Sibley and colleagues recently used a similar strategy to demonstrate that the molecular target for cytochalasin D in *Toxoplasma gondii* is in fact actin (Dobrowolski and Sibley, 1996).

Target validation

Once a potential target protein has been identified, it will be important to demonstrate that the binding is specific and to develop independent evidence that the identified target is the relevant one *in vivo*. This can be a difficult challenge. Transcriptional profiling is useful in some cases (Marton *et al.*, 1998). If the functional assay can be carried out in cell extracts, immuno/affinity depletion of the suspected target molecule is informative provided that activity can be reconstituted when the depleted molecule

is added back (e.g. Rosania *et al.*, 1999). An alternative approach, in cases where the small molecule binds to its putative target *in vitro*, is to determine the structure of the protein–ligand complex and test whether replacing the wild-type allele of the target gene with a copy containing engineered mutations in the determined ligand-binding site (mutations which abrogate small molecule binding *in vitro*) confers resistance to the small molecule (Eyers *et al.*, 1999; Davies *et al.*, 2000). An elegant variation on this approach is to design a modified small molecule that is inactive, then engineer a modified target that can accept (and be inhibited by) the modified small molecule. By replacing the wild-type allele of the target with the modified allele, a completely specific protein-ligand pair will have been created with which to analyse the target protein's function (Shah *et al.*, 1997; Kapoor and Mitchison, 1999; Bishop *et al.*, 2000).

Target identification: general comments

No single target identification strategy will work for all small molecules in all systems, and the optimal way to proceed is to undertake several different approaches in parallel. With the exception of guess-and-test and genetic strategies to target identification, small molecules identified as active in the primary screen should generally be considered lead compounds that will require further structural modification before they will be useful for target identification. These modifications may be directed towards increasing the small molecule's affinity for its target [which often leads to an increase in specificity (Eaton *et al.*, 1995)], or may involve the addition of affinity handles, cross-linking groups, or tags. These considerations underscore the critical role that synthetic chemistry plays in target identification. It is also clear that the more structure-activity information that can be collected in the initial screen, the better. Finally, it should be noted that all of the strategies outlined above (with the exception of the guess-and-test and genetic strategies) assume the target of the small molecule is a protein. Systematic methods to identify non-protein targets of bioactive small molecules (e.g. lipids, nucleic acids) remain to be developed.

Small molecule approaches in cellular microbiology

Over the past half century, many different collections of synthetic small molecules and natural products have been tested for antimicrobial activity. The collections have been screened for drugs that kill or arrest the growth of the test microbe, or for ones that interfere with the activity of specific microbial targets (e.g. Blondelle and Houghten, 1996; Fernandes *et al.*, 1999; Selzer *et al.*, 2000; Werbovetz, 2000). Screens for small molecules that either elicit or inhibit a particular microbiological phenotype (other than death!) have been much rarer (Rose *et al.*, 1996). We believe such phenotype-based screens have great potential for increasing our understanding of the mechanisms underlying important microbial processes and for identifying proteins involved in those processes. This is particularly true in systems where standard genetic tools are poorly developed, but even in systems with well established genetics, small molecules may be the most productive way to study essential genes or dynamic processes.

Our own ongoing efforts to use small molecules in the study of host cell invasion by *Toxoplasma gondii* illustrate some of these points. *T. gondii* is an obligate intracellular parasite. Despite the importance of invasion to both the life cycle of the parasite and the pathology of toxoplasmosis (Remington *et al.*, 1995; Black and Boothroyd, 2000), relatively little is known about the parasite proteins that mediate invasion. Many cytoskeletal, secretory, and surface proteins of the parasite have been identified but establishing a function for any one of these proteins in invasion is difficult. This is due, at least in part, to the fact that disruption of a gene essential for invasion in a haploid, obligate intracellular parasite such as *Toxoplasma* is by definition lethal (e.g. Hehl *et al.*, 2000; Rabenau *et al.*, 2001). Forward genetic screens for mediators of invasion suffer from the same problem: the most interesting of the mutants one might generate are likely to be non-viable.

Small molecules represent one way around this problem, and we have therefore undertaken a phenotype-based small molecule approach directed at identifying *Toxoplasma* proteins that function in invasion. Our early results are encouraging (N. J. Westwood *et al.* in preparation) and suggest that useful new probes for dissecting the molecular mechanisms underlying *Toxoplasma* invasion will be generated. We hope that this will be the first in a series of such approaches that will ultimately impact on other experimentally difficult aspects of *Toxoplasma* cell biology. For example, phenotype-based approaches could be used to dissect the signalling pathways that underlie tachyzoite-to-bradyzoite stage conversion (Bohne *et al.*, 1996; Soete and Dubremetz, 1996). Target-based approaches could be used to address the biological functions of the apicoplast, an essential, plastid-like organelle that is ancestrally derived from a green alga (Kohler *et al.*, 1997; McFadden and Roos, 1999).

One can imagine many other problems in cellular microbiology and host–pathogen interaction where small molecule approaches might also be useful. For example, small molecules could be used to study the mechanism by which vacuoles containing intracellular *Chlamydia* circumvent normal endosomal trafficking (Heinzen *et al.*, 1996; Taraska *et al.*, 1996; Al-Younes *et al.*, 1999). Genetic tools to address this question are currently

lacking, but one could readily imagine an automated, microscope-based screen in which cells infected with *Chlamydia* are fixed, permeabilized, and analysed by dual immunofluorescence with antibodies against a bacterial surface antigen and a lysosomal marker. Lysosomes do not normally fuse with the *Chlamydia*-containing vacuole; small molecules which cause fusion (i.e. result in co-localization of the two markers) could be used to identify the bacterial or host cell factor(s) that normally prevent this fusion from occurring. A second example would be to use small molecules to study the mechanisms underlying the remarkable process of polar tube discharge in microsporidian parasites (reviewed in Keohane and Weiss, 1998). Again, genetic tools are currently unavailable in microsporidia to address this question, but one could envision an immunofluorescence-based (Keohane *et al.*, 1999) high-throughput assay to identify small molecules capable of either eliciting polar tube discharge, or inhibiting the discharge induced by calcium ionophore (Pleshinger and Weidner, 1985). A third example would be to use small molecules to study the mechanisms of intra- and intercellular movement in the spotted fever group of *Rickettsia*. Significant differences exist between the actin-based motility of *Rickettsia* and the more extensively studied actin-based mechanisms of *Listeria* and *Shigella* (Gouin *et al.*, 1999; Heinzen *et al.*, 1999) and the rickettsial components which direct the process remain completely unknown. Studies of *Rickettsia* motility are difficult both because genetic tools are lacking and because of its obligate intracellular nature. Small molecule screening for inhibitors of *Rickettsia* motility, assayed either in cultured cells or *Xenopus* extracts (Gouin *et al.*, 1999), offers a way around these constraints.

Small molecules are powerful experimental tools for elucidating basic cell biological mechanisms. Applying target- and phenotype-based small molecule approaches to specific questions in cellular microbiology will enable otherwise intractable questions about mechanism to be addressed. At the same time, these approaches will generate novel small molecule probes that perturb particular cell biological processes (probes that may prove to be of use in other experimental systems), and they will very likely open up new possibilities for antimicrobial drug development.

Acknowledgements

The authors thank Rebecca Ward, Mary Tierney, Douglas Johnson, Mikael Elofsson, Ambrose Cheung, Rebekah DeVinney, Tim Mitchison and members of our laboratories for critical comments on this manuscript, and Henry Pelish for the figure upon which Fig. 3 is based. N.J.W. is a Royal Society University Research Fellow, and G.E.W. is the recipient of a Burroughs Wellcome Fund New Investigator Award in Molecular Parasitology.

References

Adam, D. (2001) Database of molecular probes set to boost molecular genetics. *Nature* **411:** 873.

Affleck, R.L. (2001) Solutions for library encoding to create collections of discrete compounds. *Curr Opin Chem Biol* **5:** 257–263.

Al-Obeidi, F., Hruby, V.J., and Sawyer, T.K. (1998) Peptide and peptidomimetic libraries. Molecular diversity and drug design. *Mol Biotechnol* **9:** 205–223.

Al-Younes, H.M., Rudel, T., and Meyer, T.F. (1999) Characterization and intracellular trafficking pattern of vacuoles containing *Chlamydia pneumoniae* in human epithelial cells. *Cell Microbiol* **1:** 237–247.

Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y. *et al.* (2000) Unnatural ligands for engineered proteins: new tools for chemical genetics. *Annu Rev Biophys Biomol Struct* **29:** 577–606.

Black, M.W., and Boothroyd, J.C. (2000) Lytic cycle of *Toxoplasma gondii*. *Microbiol Mol Biol Rev* **64:** 607–623.

Blake, R.A. (2001) Cellular screening assays using fluorescence microscopy. *Curr Opin Pharmacol* **1:** 533–539.

Blondelle, S.E., and Houghten, R.A. (1996) Novel antimicrobial compounds identified using synthetic combinatorial library technology. *Trends Biotechnol* **14:** 60–65.

Bohne, W., Parmley, S.F., Yang, S., and Gross, U. (1996) Bradyzoite-specific genes. *Curr Top Microbiol Immunol* **219:** 81–91.

Brady, S.F., Stauffer, K.J., Lumma, W.C., Smith, G.M., Ramjit, H.G., Lewis, S.D. *et al.* (1998) Discovery and development of the novel potent orally active thrombin inhibitor *N*-(9-hydroxy-9-fluorenecarboxy) prolyl trans-4-aminocyclohexylmethyl amide (L-372,460): coapplication of structure-based design and rapid multiple analogue synthesis on solid support. *J Med Chem* **41:** 401–406.

Brush, M. (1999) Automated Laboratories. *Scientist* **13** (4): 22.

Cameron, L.A., Giardini, P.A., Soo, F.S., and Theriot, J.A. (2000) Secrets of actin-based motility revealed by a bacterial pathogen. *Nat Rev Mol Cell Biol* **1:** 110–119.

Celli, J., Deng, W., and Finlay, B.B. (2000) Enteropathogenic *Escherichia coli* (EPEC) attachment to epithelial cells: exploiting the host cell cytoskeleton from the outside. *Cell Microbiol* **2:** 1–9.

Chardin, P., and McCormick, F. (1999) Brefeldin A: the advantage of being uncompetitive. *Cell* **97:** 153–155.

Cheung, A., Westwood, N., Chen, I., Mitchison, T., and Straight, A. (2001) Blebbistatin: a cell permeable inhibitor of non-muscle myosin II. *Mol Biol Cell* **12:** 271a.

Cheung, A., Dantzig, J.A., Hollingworth, S., Baylor, S.M., Goldman, Y.E., Mitchison, T.J., and Straight, A.F. (2002) A small-molecule inhibitor of skeletal muscle myosin II. *Nat Cell Biol* **4:** 83–88.

Clark, A.M. (1996) Natural products as a resource for new drugs. *Pharm Res* **13:** 1133–1144.

Cooper, J.A. (1987) Effects of cytochalasin and phalloidin on actin. *J Cell Biol* **105:** 1473–1478.

Coppens, I., Sinai, A.P., and Joiner, K.A. (2000) *Toxoplasma gondii* exploits host low-density lipoprotein receptor-mediated endocytosis for cholesterol acquisition. *J Cell Biol* **149:** 167–180.

Cortese, J.D. (2000) Well Read. *Scientist* **14** (19): 24.

Crews, C.M., and Splittgerber, U. (1999) Chemical genetics: exploring and controlling cellular processes with chemical probes. *Trends Biochem Sci* **24:** 317–320.

Davies, S.P., Reddy, H., Caivano, M., and Cohen, P. (2000)

3 Eisenberg, D. *et al.* (2000) Protein function in the post-genomics era. *Nature* 405, 823–826
4 Burley, S.K. *et al.* (1999) Structural genomics: beyond the human genome project. *Nat. Genet.* 23, 151–157
5 Hendrickson, W.A. (2000) Synchrotron crystallography. *Trends Biochem. Sci.* 25, 637–643
6 Schweitzer, B. *et al.* (2000) Inaugural article: immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. *Proc. Natl. Acad. Sci. U. S. A.* 97, 10113–10119
7 Legrain, P. *et al.* (2001) Protein–protein interaction maps: a lead towards cellular functions. *Trends Genet.* 17, 346–352
8 Ideker, T. *et al.* (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292, 929–934

Yuk Fai Leung
Chi Pui Pang*
Dept of Ophthalmology & Visual Sciences, The Chinese University of Hong Kong, Hong Kong Eye Hospital, 3/F, 147K Argyle Street, Kowloon, Hong Kong.
*e-mail: cppang@cuhk.edu.hk

Techniques & Applications

Small-molecule metabolism: an enzyme mosaic

Sarah A. Teichmann, Stuart C.G. Rison, Janet M. Thornton, Monica Riley, Julian Gough and Cyrus Chothia

***Escherichia coli* has been a popular organism for studying metabolic pathways. In an attempt to find out more about how these pathways are constructed, the enzymes were analysed by defining their protein domains. Structural assignments and sequence comparisons were used to show that 213 domain families constitute ~90% of the enzymes in the small-molecule metabolic pathways. Catalytic or cofactor-binding properties between family members are often conserved, while recognition of the main substrate with change in catalytic mechanism is only observed in a few cases of consecutive enzymes in a pathway. Recruitment of domains across pathways is very common, but there is little regularity in the pattern of domains in metabolic pathways. This is analogous to a mosaic in which a stone of a certain colour is selected to fill a position in the picture.**

According to the *Concise Oxford Dictionary*, a mosaic is 'a picture...produced by an arrangement of small variously coloured pieces of glass or stone'. A mosaic is analogous in several ways to small-molecule metabolic pathways. In particular, the enzymes that form the metabolic pathways belong to a limited set of protein families, like the set of different coloured pieces available to the artist to construct the mosaic. Furthermore, the picture of the mosaic as a whole is meaningful, even though there is no discernible repeating pattern in the way the pieces are arranged; instead, each piece has been selected to fill a space with the necessary colour to make the mosaic picture. Likewise, domains in enzymes appear to have been selected from a protein family in an unsystematic way to fill a position in a pathway for the functional features of that family.

Box 1. Determining the domain structure and family membership of enzymes

Structural domains

The domain definitions and evolutionary relationships of the proteins of known structure are described in the Structural Classification of Proteins (SCOP) database[a] (http://scop.mrc-lmb.cam.ac.uk/scop/). In SCOP, domains are structural but also evolutionary units, so a domain has to be observed on its own in a structure or combined with several different domains to be classified as a domain. The phenylalanyl-tRNA synthetase large chain is shown as an example of a multi-domain polypeptide chain (Fig. I).

Domains are classified into superfamilies on the basis of sequence, as well as structural and functional features that are shared by all the domains in a superfamily.

Gough *et al.*[b] used the domains from SCOP version 1.53 as seed sequences to build a type of profile called Hidden Markov Models. (The specific method is described by Karplus *et al.*[c]) The database of Hidden Markov Models is available at http://stash.mrc-lmb.cam.ac.uk/SUPERFAMILY/.

These models were then scanned against the *Escherichia coli* enzymes to identify domains in the enzymes. The family membership of the *E. coli* domains was inferred from the SCOP superfamily membership of the homologous SCOP domain.

Sequence domains

The regions of the *E. coli* enzymes not matched by a structural domain were compared using the multiple sequence comparison procedure PSI-BLAST[d], and then clustered into families as described by Park and Teichmann[e].

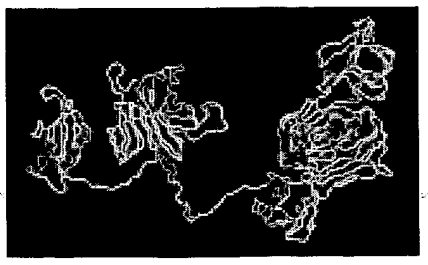


Fig. I An example of a multi-domain polypeptide chain.

References
a Murzin, A. *et al.* (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540
b Gough, J. *et al.* (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* 313, 903–919
c Karplus, K. *et al.* (1998) Hidden Markov Models for detecting remote protein homologies. *Bioinformatics* 14, 846–856
d Altschul, S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402
e Park, J. and Teichmann, S.A. (1998) DIVCLUS: an automatic method in the GEANFAMMER package that finds homologous domains in single- and multi-domain proteins. *Bioinformatics* 14, 144–150

  PII: S0167-7799(01)01813-3

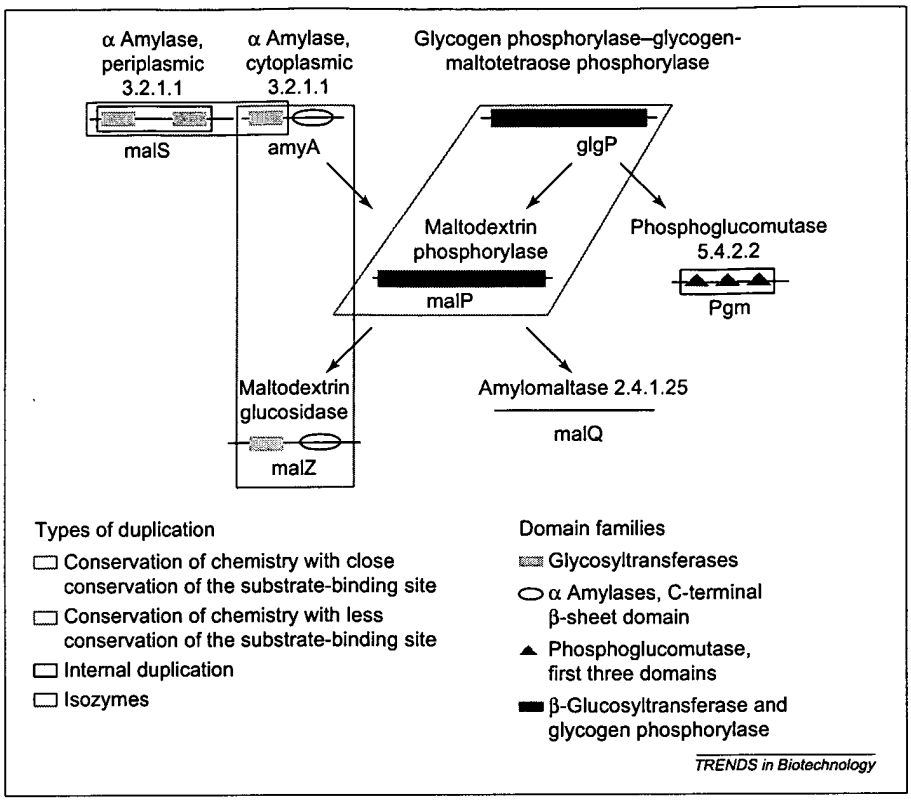


Fig. 1. Glycogen catabolism pathway. The enzymes are represented by black lines and the structural domains by coloured shapes in N-to-C-terminal order on the polypeptide chain. The arrows represent the flux of substrates and products through the pathway. There are two duplications with conservation of catalytic mechanism in this pathway. One is in consecutive enzymes (e.g. glgP and malP), therefore there is also close conservation of substrate-binding site, whereas the other duplication occurs for enzymes one step apart (e.g. amyA and malZ), with less conservation of the substrate-binding site. There are also internal duplications, in which the same type of domain occurs several times in one polypeptide sequence (malS and pgm) and isozymes (malS and amyA).

The 'colours' of the enzymes in the mosaic of *Escherichia coli* small-molecule metabolic pathways were determined by assigning the domains in each enzyme to a protein family. These protein families were derived from a combination of sequence and structural information (Box 1). Like roughly hewn mosaic pieces of one colour, the domains that belong to one family are not identical, but can be very divergent. The result of the domain assignments is a description of the structural anatomy of metabolic pathways and their enzymes, for example those involved in glycogen catabolism (Fig. 1). Such a clarification of the domain structure of enzymes provides a picture of the structural anatomy of the individual enzymes in the metabolic pathways and allows investigation into any patterns in duplicated enzyme domains within and across the metabolic pathways.

Structural anatomy of *E. coli* small-molecule metabolic enzymes

The metabolic pathways in *E. coli* are probably the most thoroughly studied of any organism. Although the details of the enzymes and metabolic pathways will differ from organism to organism, the principles of the structure and evolution of the pathways would be expected to apply across all organisms. The EcoCyc database[1] contains comprehensive information on small-molecule metabolism in *E. coli*, and the 106 pathways and the corresponding 581 enzymes described in this database were used in the present study. The results of the domain assignment procedure (shown in Box 1 and described in detail in Ref. 2) gave a total of 722 domains in 213 families in 510 (88%) of the *E. coli* small-molecule metabolism (SMM) enzymes (summarized in Box 2 and Table 1). There are, on average, 3.4 domains per family, which shows that even this basic set of pathways is the product of extensive duplication of domains within its enzymes. The distribution of family sizes of the 213 families is roughly exponential: 74 families in *E. coli* SMM have only one domain, and the largest family, the Rossmann domains, has 53 domains.

There has been not only extensive duplication of domains but also combinations of domains in these pathways, as exemplified by the fact that 722 domains are assigned to only 510 enzymes. Two-thirds of the 213 families have at least one domain that is adjacent (within 75 residues) to another assigned domain in one of the SMM proteins. Most families have only one or two types of domain partners in a fixed N-to-C-terminal orientation, but the Rossmann domain family has 12 different partner families.

Figure 2 illustrates some of the enzymes that contain Rossmann domains. Half of the SMM enzymes are single-domain proteins, similar to the dihydrobenzoate dehydrogenase (entA) in Figure 2. A quarter of all SMM enzymes contain two domains. For example, the NAD-linked malic enzyme (sfcA) shown in Fig. 2 consists of a Rossmann domain and an amino acid dehydrogenase-like domain. Of the 141 families that are adjacent to another

Box 2. Pathways, proteins, domains and families

Number of metabolic pathways	106
Number of proteins	581
Number of proteins of known sequence	569
Number of proteins with assigned domains	510
Structural domains	695 in 202 families
Sequence domains	27 in 11 families

Table 1. Numbers of domains in enzymes

Number of domains (n)	Numbers of sequences completely matched by n domains	Numbers of sequences partly matched by n domains
1	271	77
2	96	26
3	28	5
4	2	3
5	1	–
6	1	–
Total number of proteins	399	111

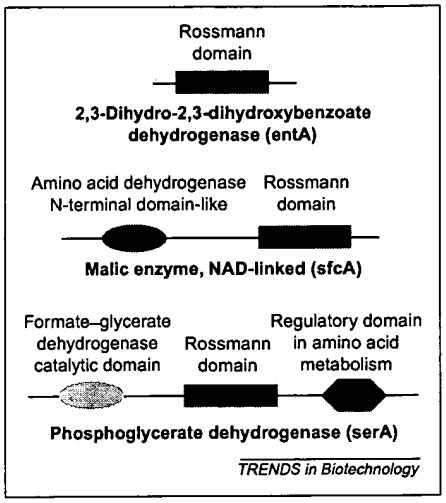


Fig. 2. Rossmann domains in enzymes. The polypeptide chains of enzymes are represented by black lines and the structural domains are represented by shapes from left to right in their N-to-C-terminal order. Examples of single-domain, two-domain and three-domain enzymes containing Rossmann domains are given, showing how domains from this family combine with other domains in different ways.

assigned domain in the SMM enzymes, 73% combine with only one type of domain. The Rossmann domain family, however, is versatile in that it can combine with more than one type of domain. Figure 2 shows two domain neighbours, in addition to those of the amino acid dehydrogenase-like family, in phosphoglycerate dehydrogenase (serA). Like the phosphoglycerate dehydrogenase (serA), a sixth of all *E. coli* SMM enzymes contain three to six domains. Half of the SMM enzymes are multi-domain enzymes, and almost three-quarters of the domain families in these enzymes have at least one domain member that is adjacent to another assigned domain in one of the SMM enzymes.

It is clear that even proteins as fundamental to the functioning of a free-living cell, and also as ancient as the central SMM enzymes, are not all simple single-domain enzymes but are the product of extensive domain combinations. Therefore, either SMM enzymes developed by fusions and recombinations from a more basic set of proteins, which were single-domain proteins, or combinations of two or more domains occurred first, and then domains later split and recombined to crystallize as individual evolutionary units, the domains that are recognized today.

Evolution of *E. coli* small-molecule metabolic pathways

Information about the domain structures of the individual enzymes can be used to investigate aspects of the evolution of metabolic pathways. Of the 213 domain families, 144 have members distributed across different pathways. The 69 families that are active in only one pathway are all small: 67 have one or two members, one has three members and one has four members. This distribution shows that the evolution of metabolic pathways involved widespread recruitment of enzymes to different pathways, which supports Jensen's model of pathway evolution[3].

Types of conservation of domain duplications

It is helpful when discussing pathway evolution to distinguish between different types of duplications of enzymes and their domains. Figure 1 shows multiple copies of the four types of domains in the glycogen catabolism pathway. The glycosyltransferase domain family (yellow) and the phosphoglucomutase domains (green) recur within the individual proteins periplasmic α amylase (malS) and phosphoglucomutase (pgm). This type of duplication is termed internal duplication and can only take place within pathways. Duplication of domains in enzymes that are isozymes can also only occur within pathways. Glycosyltransferase domains are also present in periplasmic (malS) and cytoplasmic α amylase (amyA) and in the maltodextrin glucosidase (malZ). The duplication between α amylase and maltodextrin glucosidase conserves catalytic mechanism because enzymes hydrolyse glucosidic linkages. Similarly, the two phosphorylase domains (shown in blue) conserve reaction chemistry because both glycogen phosphorylase (glgP) and maltodextrin phosphorylase (malP) are phosphorylases acting on different substrates. Recent studies have described this evolutionary mechanism in detail and show how mutations in active site residues produce new catalytic properties for enzymes[4–7]. There are two further types of duplication that do not occur in the glycogen catabolism pathway: duplication of cofactor- or minor substrate-binding domains such as Rossmann domains and duplication with conservation of the substrate-binding site but change in catalytic mechanism.

Duplications within pathways

Of the different types of duplication listed previously, internal duplication and duplication that occurs in isozymes are frequent within pathways. Duplication with conservation of a cofactor- or minor substrate-binding site is also frequent within pathways. Within the entire set of almost 600 enzymes, there are only six examples of duplications in pathways with conservation of the major substrate-binding site and a change in the catalytic mechanism (Table 2). This means that duplications in pathways are driven by similarity in catalytic mechanism much more than by similarity in the substrate-binding pocket. This disagrees with Horowitz' model of retrograde evolution[8], in which it is suggested that enzymes within a pathway are related to each other. In fact, more enzymes that are separated by one catalytic step share a domain (11%) than do consecutive

Table 2. Conservation of the main substrate-binding site with change in reaction catalysed within a pathway.

Superfamily and pathway	Enzymes
Phosphoenolpyruvate and pyruvate $(\alpha/\beta)_8$ barrels in fermentation	pykF/pykA, ppc
Ribulose-phosphate binding $(\alpha/\beta)_8$ barrels in tryptophan biosynthesis	trpA, trpC
[a]P-binding α/β barrels in histidine, purine and pyrimidine biosynthesis	hisA, hisF
Phosphoribosyltransferases (PRTases) in histidine, purine and pyrimidine biosynthesis	prsA,purF and prsA, pyrE
dUTPase domains in deoxypyrimidine nucleotide/nucleoside metabolism	dcd, dut
Inosine monophosphate dehydrogenase $(\alpha/\beta)_8$ barrels in nucleotide metabolism	guaB, guaC

[a]The P-binding α/β barrels are a diverse family of α/β barrels that are likely to be related because they share a phosphate-binding site in the loop between β-strand 7 and α-helix 7 and the N-terminus of an additional helix 8'.

These examples are the only detected cases of enzymes that belong to the same family and share a similar binding site for the main substrate within a pathway, but change their reaction chemistry. Therefore, this type of conservation is much more rare than change in substrate specificity with conservation of chemistry in metabolic pathways.

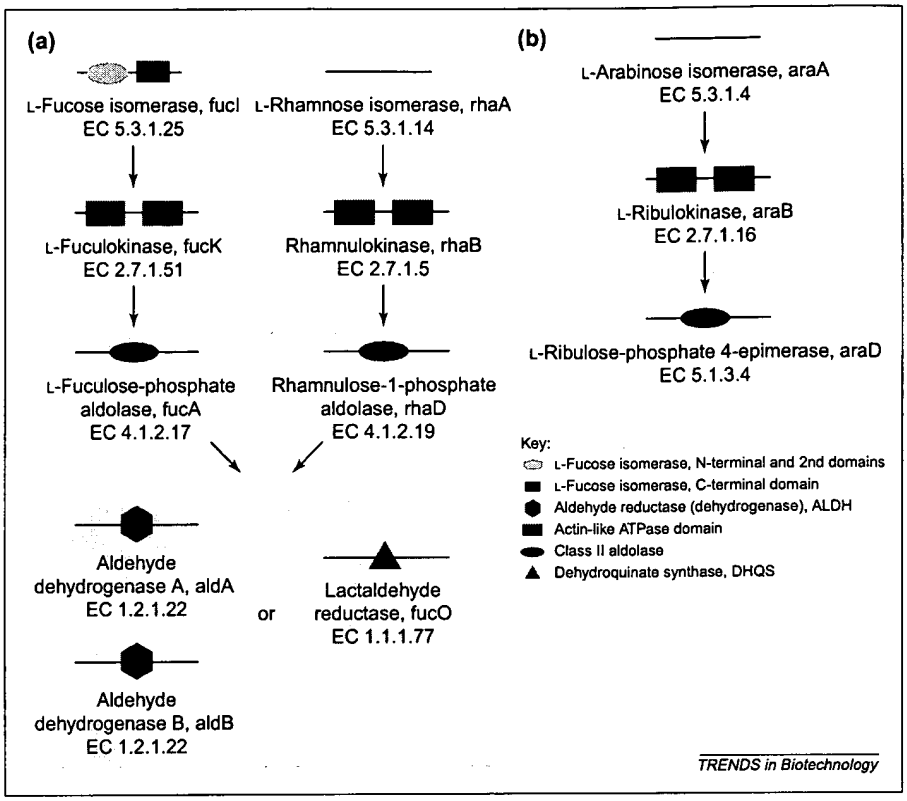


Fig. 3. Fucose, rhamnose and L-arabinose catabolism. (a) Fucose and rhamnose. A superpathway exists in EcoCyc that consists of the fucose and rhamnose catabolism subpathways. An example of serial recruitment and of 'parallel' enzymes is shown (boxed). Serial recruitment has occurred because fucK (L-fuculokinase) is homologous to rhaB (rhamnulokinase), and fucA (L-fuculose-phosphate aldolase) is homologous to rhaD (rhamnulose-1-phosphate aldolase). fucA and rhaD have the same product, and are both followed by aldA or aldB or fucO, and are thus parallel enzymes. The enzyme classification (EC) numbers for each enzyme are given. (b) L-Arabinose. AraB is homologous to fucK and rhaB in (a), and araD is homologous to fucA and rhaD in (a). The three pairs of enzymes are an example of serial recruitment, as supported by their similar positions on the *Escherichia coli* chromosome: the genes in each pair are divided by one gene on the chromosome.

enzymes (6%), indicating that there is no bias for duplication between enzymes that are close to each other in a pathway.

Duplications within pathways occur relatively frequently in situations such as that shown in Fig. 3a, in which L-fuculose-phosphate aldolase (fucA) and rhamnulose-1-phosphate aldolase (rhaD) are homologous. In this type of case, two enzymes are followed by the same enzyme(s) in a pathway and hence have the same or similar products. Alternatively, two enzymes can also be 'parallel' when both have the same precursor enzyme in a pathway and thus have the same or similar substrates. 13% (same or similar substrates) and 17% (same or similar products) of these parallel enzymes in pathways have homologous domains.

Of the eight cases in which two enzymes are followed by the same enzyme, as in fucose and rhamnose catabolism, there are two cases, such as L-fuculose-phosphate aldolase (fucA) and rhamnulose-1-phosphate aldolase (rhaD), in which the two enzymes catalyse similar reactions and have the same product. In all the other cases, the products are merely similar, so that the enzyme that follows in the pathway possesses multiple substrate specificity. In five of the seven cases where two enzymes act on the same substrate, the two enzymes carry out similar reactions, often using a different second substrate in a reaction, such as a transferase or synthase reaction.

Duplications across pathways

As mentioned, all the larger domain families in the metabolic pathways have members in more than one pathway, thus duplications across pathways are extremely common. However, it appears that little of this recruitment takes place in an ordered fashion. Examples of serial recruitment, where two enzymes in one pathway are recruited to another pathway in the same order, such as L-fuculokinase (fucK) and L-fuculose-phosphate aldolase (fucA), rhamnulokinase (rhaB) and rhamnulose-1-phosphate aldolase (rhaD), and L-ribulokinase (araB) and L-ribulose phosphate 4-epimerase (araD) in Fig. 3, are very rare. If duplication of large portions of the bacterial chromosome takes place, and all the genes in a duplicated portion were used to form a new pathway, serial recruitment would be expected. In fact, only 89 out of 26 341 (0.3%) possible pairs of enzymes are homologous in both the first and second enzymes. Only seven of these 89 pairs of doublets of enzymes have the genes for both doublets close to each other on the chromosome, which suggests that the two initial enzymes might have been duplicated as one portion. The three kinase– and aldolase–epimerase pairs of enzymes involved in sugar catabolism are a good example of this rare situation: all three pairs are one gene apart on the *E. coli* chromosome.

Conclusions and discussion

This description of how a relatively small repertoire of 213 domain families constitutes 90% of the enzymes in the *E. coli* small-molecule metabolic pathways is, to some extent, paradoxical. Although the SMM enzymes have arisen by extensive duplication, with an average of 3.4 domain members per SMM family, the distribution of families within and across pathways is complex: there is little repetition of domains in consecutive steps of pathways and little serial homology across pathways. Together with the analysis of the chromosomal locations of genes, it is evident that metabolic pathways have, in general, not arisen by duplication of large portions of the *E. coli* chromosome, either to extend a pathway or to make a new pathway. There are a few well known exceptions to this, such as the enzymes involved in the fucose, rhamnose and arabinose catabolic pathways. Similarly, duplication of enzymes that conserve a substrate-binding site is rare, otherwise the fraction of consecutive homologous enzymes would be larger. The main pressure for selection for enzymes in pathways appears to be either their catalytic mechanism or cofactor-binding properties. This pattern of evolution has resulted in a mosaic of enzyme domains optimized for smooth-functioning

small-molecule metabolism in *E. coli*, with little order in the pattern of domains with respect to position within or between pathways.

Selection based entirely on function, and specifically reaction chemistry, was termed 'patchwork evolution' by Lazcano and Miller and also by Copley in a discussion of the pathway for the degradation of pentachlorophenol by the soil micro-organism *Sphingomonas chlorophenolica*[10]. Pentachlorophenol was introduced into the environment in 1936, and is not produced naturally, so it is probable that the pathway evolved in the past few decades. The pathway involves three enzymes, which were recruited in a 'patchwork' manner from the enzymes that break down naturally occurring chlorinated phenols.

Recently, recruitment of enzymes across metabolic pathways was observed in a study of the distribution of $(\alpha/\beta)_8$ barrels by Copley and Bork[11], and in a review on structural genomics of metabolic pathways by Erlandsen and colleagues[12]. The comprehensive structural assignments to 90% of the enzymes in all *E. coli* small-molecule metabolic pathways described in the present article confirm that pathways are constructed by recruitment on the basis of catalytic mechanism, with few instances of duplication of enzymes within a pathway or serial recruitment across pathways.

Acknowledgements

SAT has a Beit Memorial Fellowship. SCGR acknowledges support from GlaxoSmithkline and MR acknowledges support from the NIH and the NASA Astrobiology Institute. The authors acknowledge computational support from the BBSRC.

References

1 Karp, P. *et al.* (1999) EcoCyc: electronic encyclopedia of *E. coli* genes and metabolism. *Nucleic Acids Res.* 27, 55–58
2 Teichmann, S.A. *et al.* (2001) The Evolution and Structural Anatomy of the Small Molecule Metabolic Pathways in *Escherichia coli*. *J. Mol. Biol.* 311, 693–708
3 Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* 30, 409–425
4 Murzin, A.G. (1993) Can homologous proteins evolve different enzymatic activities? *Trends Biochem. Sci.* 18, 403–405
5 Neidhart, D.J. and Petsko, G. (1990) Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. *Nature* 347, 692–694
6 Babbitt, P.C. and Gerlt, J.A. (1997) Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities. *J. Biol. Chem.* 272, 30591–30594
7 Petsko, G.A. *et al.* (1993) On the origin of enzymatic species. *Trends. Biochem. Sci.* 18, 372–376
8 Horowitz, N.H. (1945) On the evolution of biochemical syntheses. *Proc. Natl. Acad. Sci. U. S. A.* 31, 152–157
9 Lazcano, A. and Miller, S.C. (1999) One the origin of metabolic pathways. *J. Mol. Evol.* 49, 424–431
10 Copley, S.D. (2000) Evolution of a metabolic pathway for degradation of a toxic xenobiotic: the pathwork approach. *Trends Biochem. Sci.* 25, 261–265
11 Copley, R.R. and Bork, P. (2000) Homology of $(\beta\alpha)_8$ barrels: implications for the evolution of metabolic pathways. *J. Mol. Biol.* 303, 627–641
12 Erlandsen, H. *et al.* (2000) Combining structural genomics and enzymology: completing the picture in metabolic pathways and enzyme active sites. *Curr. Opin. Struct. Biol.* 10, 719–730

Sarah A. Teichmann*
Stuart C. G. Rison
Janet M. Thornton
Dept of Biochemistry and Molecular Biology, University College London, Darwin Building, Gower Street, London, UK WC1E 6BT.
*e-mail: sat@mrc-lmb.cam.ac.uk

Monica Riley
Josephine Bay Paul Centre for Comparative Molecular Biology and Evolution, 7 MBL St, Woods Hole, MA 02543-1015, USA.

Julian Gough
Cyrus Chothia
MRC Laboratory of Molecular Biology, Hills Road, Cambridge, UK CB2 1TQ.

Center for Molecular Biodiversity and Evolution set up

The biotechnology company BRAIN (Zwingenberg, Germany) and scientists from the Institute of Genetics and Microbiology (Technical University, Darmstadt, Germany) have jointly set up the Center for Molecular Biodiversity and Evolution (ZEB) at the Technical University of Darmstadt. The Center was set up with the aim of exploring the >99% of microorganisms in a typical soil sample that cannot be cultivated and to search for new enzymes and bioactive molecules. The Center's main goal is to isolate the collective genomes of a microbial community, the 'metagenome', by directly isolating DNA from soil and incorporating it into BioArchives (recombinant DNA libraries containing environmental DNA). The ZEB will be headed by Christa Schelper and represents a promising cooperation between academia and industry.

First results of collaboration between Graffinity and Aventis announced

Graffinity Pharmaceuticals (Heidelberg, Germany) has recently announced the first results in its chemical microarray collaboration with Aventis Pharma (Frankfurt, Germany). Graffinity uses chemical genomics to convert lead targets into small-molecule pharmaceuticals. The agreement between Graffinity and Aventis was first announced in May 2001 – Graffinity was to synthesise exclusive arrays for Aventis to discover novel drug leads.

Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem J* **351:** 95–105.

Degterev, A., Lugovskoy, A., Cardone, M., Mulley, B., Wagner, G., Mitchison, T., and Yuan, J. (2001) Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. *Nat Cell Biol* **3:** 173–182.

Delneri, D., Brancia, F.L., and Oliver, S.G. (2001) Towards a truly integrative biology through the functional genomics of yeast. *Curr Opin Biotechnol* **12:** 87–91.

Dobrowolski, J.M., and Sibley, L.D. (1996) *Toxoplasma* invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell* **84:** 933–939.

Dolle, R.E. (2001) Comprehensive survey of combinatorial library synthesis: 2000. *J Comb Chem* **3:** 477–517.

Driever, W., Solnica-Krezel, L., Schier, A.F., Neuhauss, S.C., Malicki, J., Stemple, D.L. *et al.* (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123:** 37–46.

Eaton, B.E., Gold, L., and Zichi, D.A. (1995) Let's get specific: the relationship between specificity and affinity. *Chem Biol* **2:** 633–638.

Eyers, P.A., van den Ijssel, P., Quinlan, R.A., Goedert, M., and Cohen, P. (1999) Use of a drug-resistant mutant of stress-activated protein kinase 2a/p38 to validate the *in vivo* specificity of SB 203580. *FEBS Lett* **451:** 191–196.

Feng, Y., Jadhav, A., Roth, M., Melancon, P., Mitchison, T., and Kirchhausen, T. (2001) Ergostatin, a Golgi disrupting compound, targets ARF GTPase. *Mol Biol Cell* **12:** 272a.

Fernandes, P.B., Menzel, R., Hardy, D.J., Tse-Dinh, Y.C., Warren, A., and Elsemore, D.A. (1999) Microbial resistance: novel screens for a contemporary problem. *Med Res Rev* **19:** 559–568.

Foster, B.A., Coffey, H.A., Morin, M.J., and Rastinejad, F. (1999) Pharmacological rescue of mutant p53 conformation and function. *Science* **286:** 2507–2510.

Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. (1991) General method for rapid synthesis of multicomponent peptide mixtures. *Int J Pept Protein Res* **37:** 487–493.

Gouin, E., Gantelet, H., Egile, C., Lasa, I., Ohayon, H., Villiers, V. *et al.* (1999) A comparative study of the actin-based motilities of the pathogenic bacteria *Listeria monocytogenes, Shigella flexneri* and *Rickettsia conorii*. *J Cell Sci* **112:** 1697–1708.

Grabley, S., and Thiericke, R. (1999) Bioactive agents from natural sources: trends in discovery and application. *Adv Biochem Eng Biotechnol* **64:** 101–154.

Gray, N.S., Wodicka, L., Thunnissen, A.M., Norman, T.C., Kwon, S., Espinoza, F.H. *et al.* (1998) Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* **281:** 533–538.

Grozinger, C.M., Chao, E.D., Blackwell, H.E., Moazed, D., and Schreiber, S.L. (2001) Identification of a class of small molecule inhibitors of the sirtuin family of NAD-dependent deacetylases by phenotypic screening. *J Biol Chem* **276:** 38837–38843.

Gura, T. (2000) A Chemistry Set for Life. *Nature* **407:** 282–284.

Hackstadt, T., Rockey, D.D., Heinzen, R.A., and Scidmore, M.A. (1996) *Chlamydia trachomatis* interrupts an exocytic pathway to acquire endogenously synthesized sphingomyelin in transit from the Golgi apparatus to the plasma membrane. *EMBO J* **15:** 964–977.

Haffter, P., Granato, M., Brand, M., Mullins, M.C., Hammerschmidt, M., Kane, D.A. *et al.* (1996) The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio Development* **123:** 1–36.

Hartwell, L., Culotti, J., Pringle, J.R., and Reid, B.J. (1974) Genetic control of the cell division cycle in yeast. *Science* **183:** 45–51.

Hehl, A.B., Lekutis, C., Grigg, M.E., Bradley, P.J., Dubremetz, J.F., Ortega-Barria, E., and Boothroyd, J.C. (2000) *Toxoplasma gondii* homologue of *Plasmodium* Apical Membrane Antigen 1 is involved in invasion of host cells. *Infect Immun* **68:** 7078–7086.

Heinzen, R.A., Scidmore, M.A., Rockey, D.D., and Hackstadt, T. (1996) Differential interaction with endocytic and exocytic pathways distinguish parasitophorous vacuoles of *Coxiella burnetii* and *Chlamydia trachomatis*. *Infect Immun* **64:** 796–809.

Heinzen, R.A., Grieshaber, S.S., Van Kirk, L.S., and Devin, C.J. (1999) Dynamics of actin-based movement by *Rickettsia rickettsii* in vero cells. *Infect Immun* **67:** 4201–4207.

Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D. *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell* **102:** 109–126.

Jackson, C.L. (2000) Brefeldin A revealing the fundamental principles governing membrane dynamics and protein transport. *Subcell Biochem* **34:** 233–272.

Kapoor, T.M., Mayer, T.U., Coughlin, M.L., and Mitchison, T.J. (2000) Probing spindle assembly mechanisms with monastrol, a small molecule inhibitor of the mitotic kinesin, Eg5. *J Cell Biol* **150:** 975–988.

Kapoor, T.M., and Mitchison, T.J. (1999) Allele-specific activators and inhibitors for kinesin. *Proc Natl Acad Sci USA* **96:** 9106–9111.

Kariv, I., and Chung, D.Y. (2001) Novel HTS technologies: development of miniaturized assay formats. *Current Drug Discovery* April: 28–34.

Kassel, K.B. (2001) Combinatorial chemistry and mass spectrometry in the 21st century drug discovery laboratory. *Chem Rev* **101:** 255–267.

Kenny, B., and Finlay, B.B. (1997) Intimin-dependent binding of enteropathogenic *Escherichia coli* to host cells triggers novel signaling events, including tyrosine phosphorylation of phospholipase C-gamma1. *Infect Immun* **65:** 2528–2536.

Kenny, B.A., Bushfield, M., Parry-Smith, D.J., Fogarty, S., and Treherne, J.M. (1998) The application of high-throughput screening to novel lead discovery. *Prog Drug Res* **51:** 245–269.

Keohane, E.M., and Weiss, L.M. (1998) Characterization and function of the microsporidian polar tube: a review. *Folia Parasitol* **45:** 117–127.

Keohane, E.M., Orr, G.A., Takvorian, P.M., Cali, A., Tanowitz, H.B., Wittner, M., and Weiss, L.M. (1999) Polar tube proteins of microsporidia of the family encephalitozoonidae. *J Eukaryot Microbiol* **4:** 1–5.

Kim, S.K. (2001) Http://*C.elegans*: mining the functional genomic landscape. *Nat Rev Genet* **2:** 681–689.

King, R.W. (1999) Chemistry or biology: which comes first after the genome is sequenced? *Chem Biol* **6:** R327–R333.

Klebe, G. (2000) Recent developments in structure-based drug design. *J Mol Med* **78:** 269–281.

Knockaert, M., Gray, N., Damiens, E., Chang, Y.T., Grellier, P., Grant, K. *et al.* (2000) Intracellular targets of cyclin-dependent kinase inhibitors: identification by affinity chromatography using immobilized inhibitors. *Chem Biol* **7:** 411–422.

Knockaert, M., Wieking, K., Schmitt, S., Leost, M., Grant, K.M., Mottram, J.C. *et al.* (2002) Intracellular targets of paullones:

Identification following affinity purification on immobilized inhibitor. *J Biol Chem* (in press).
Kohler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J. *et al.* (1997) A plastid of probable green algal origin in Apicomplexan parasites. *Science* **275:** 1485–1489.
Lam, K.S., Lebl, M., and Krchnak, V. (1997) The 'One-Bead-One-Compound' combinatorial library method. *Chem Rev* **97:** 411–448.
MacBeath, G., and Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* **289:** 1760–1763.
Martin, Y.C. (2001) Diverse viewpoints on computational aspects of molecular diversity. *J Comb Chem* **3:** 231–250.
Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J. *et al.* (1998) Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat Med* **4:** 1293–1301.
Mattheakis, L.C., and Savchenko, A. (2001) Assay technologies for screening ion channel targets. *Curr Opin Drug Discov Devel* **4:** 124–134.
Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., and Mitchison, T.J. (1999) Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* **286:** 971–974.
Mayer, T.U., Perlman, Z.E., and Mitchison, T.J. (2001) Small Molecule Approach to Dissect Centrosome Duplication in Mammalian Cells. *Mol Biol Cell* **12:** 174a.
McDowell, W., and Schwarz, R.T. (1988) Dissecting glycoprotein biosynthesis by the use of specific inhibitors. *Biochimie* **70:** 1535–1549.
McFadden, G.I., and Roos, D.S. (1999) Apicomplexan plastids as drug targets. *Trends Microbiol* **7:** 328–333.
Meng, L., Kwok, B.H., Sin, N., and Crews, C.M. (1999) Eponemycin exerts its antitumor effect through the inhibition of proteasome function. *Cancer Res* **59:** 2798–2801.
Mitchison, T. (1994) Towards a Pharmacological Genetics. *Chem Biol* **1:** 3–6.
Morrissette, N.S., and Sibley, L.D. (2002) Cytoskeleton of apicomplexan parasites. *Microbiol Mol Biol Rev* **66:** 21–38.
Muegge, I., Heald, S.L., and Brittelli, D. (2001) Simple selection criteria for drug-like chemical matter. *J Med Chem* **44:** 1841–1846.
Nusslein-Volhard, C., and Wieschaus, E. (1980) Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287:** 795–801.
Ohizumi, Y. (1997) Application of physiologically active substances isolated from natural resources to pharmacological studies. *Jpn J Pharmacol* **73:** 263–289.
Ohlmeyer, M.H., Swanson, R.N., Dillard, L.W., Reader, J.C., Asouline, G., Kobayashi, R., Wigler, M., and Still, W.C. (1993) Complex synthetic chemical libraries indexed with molecular tags. *Proc Natl Acad Sci USA* **90:** 10922–10926.
Ohmori, H., and Toyama, S. (1992) Direct proof that the primary site of action of cytochalasin on cell motility processes is actin. *J Cell Biol* **116:** 933–941.
Pelish, H.E., Westwood, N.J., Feng, Y., Kirchhausen, T., and Shair, M.D. (2001) Use of biomimetic diversity-oriented synthesis to discover galanthamine-like molecules with biological properties beyond those of the natural product. *J Am Chem Soc* **123:** 6740–6741.
Pelkmans, L., Puntener, D., and Helenius, A. (2002) Local actin polymerization and dynamin recruitment in SV40-induced internalization of caveolae. *Science* **296:** 535–539.
Peterson, R.T., Link, B.A., Dowling, J.E., and Schreiber, S.L. (2000) Small molecule developmental screens reveal the logic and timing of vertebrate development. *Proc Natl Acad Sci USA* **97:** 12965–12969.
Peterson, J., Lokey, R.S., Mitchison, T., and Kirschner, M.W. (2001) A chemical inhibitor of N-WASP reveals a new mechanism for targeting protein interactions. *Proc Natl Acad Sci USA* **98:** 10624–10629.
Peyroche, A., Antonny, B., Robineau, S., Acker, J., Cherfils, J., and Jackson, C.L. (1999) Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: Involvement of specific residues of Sec7 domain. *Mol Cell* **3:** 275–285.
Pleshinger, J., and Weidner, E. (1985) The microsporidian spore invasion tube. IV. Discharge activation begins with pH-triggered Ca2+ influx. *J Cell Biol* **100:** 1834–1838.
Quillan, J.M., Jayawickreme, C.K., and Lerner, M.R. (1995) Combinatorial diffusion assay used to identify topically active melanocyte-stimulating hormone receptor antagonists. *Proc Natl Acad Sci USA* **92:** 2894–2898.
Rabenau, K.E., Sohrabi, A., Tripathy, A., Reitter, C., Ajioka, J.W., Tomley, F.M. *et al.* (2001) TgM2AP participates in *Toxoplasma gondii* invasion of host cells and is tightly associated with the adhesive protein TgMIC2. *Mol Microbiol* **41:** 537–547.
Rademann, J., and Jung, G. (2000) Techview: drug discovery. Integrating combinatorial synthesis and bioassays. *Science* **287:** 1947–1948.
Remington, J.S., McLeod, R., and Desmonts, G. (1995) Toxoplasmosis. In: *Infectious Diseases of the Fetus and the Newborn Infant.* Remington, J.S., and Klein, J.O. (eds). Philadelphia: W.B. Saunders Company, pp. 140–267.
Rodriguez, A., Rioult, M.G., Ora, A., and Andrews, N.W. (1995) A trypanosome-soluble factor induces IP3 formation, intracellular Ca2+ mobilization and microfilament rearrangement in host cells. *J Cell Biol* **129:** 1263–1273.
Rosania, G.R., Merlie, J. Jr, Gray, N., Chang, Y.T., Schultz, P.G., and Heald, R. (1999) A cyclin-dependent kinase inhibitor inducing cancer cell differentiation: biochemical identification using *Xenopus* egg extracts. *Proc Natl Acad Sci USA* **96:** 4797–4802.
Rosania, G.R., Chang, Y.T., Perez, O., Sutherlin, D., Dong, H., Lockhart, D.J., and Schultz, P.G. (2000) Myoseverin, a microtubule-binding molecule with novel cellular effects. *Nat Biotechnol* **18:** 304–308.
Rose, P., Gorman, J., Kurtz, S., Pramathesh, P., and Prabhavathi, F. (1996) The successful partnership of biotechnology based screen development with high throughput screening. *Network Sci* **2:** 1–12.
Ryning, F.W., and Remington, J.S. (1978) Effect of cytochalasin D on *Toxoplasma gondii* cell entry. *Infect Immun* **20:** 739–743.
Schreiber, S.L. (1998) Chemical genetics resulting from a passion for synthetic organic chemistry. *Bioorg Med Chem* **6:** 1127–1152.
Schreiber, S.L., and Crabtree, G.R. (1992) The mechanism of action of cyclosporinA and FK506. *Immunol Today* **13:** 136–142.
Schwartzman, J.D., and Pfefferkorn, E.R. (1983) Immunofluorescent localization of myosin at the anterior pole of the coccidian, *Toxoplasma gondii*. *J Protozool* **30:** 657–661.
Selzer, P.M., Brutsche, S., Wiesner, P., Schmid, P., and Mullner, H. (2000) Target-based drug discovery for the development of novel antiinfectives. *Int J Med Microbiol* **290:** 191–201.
Setti, E.L., and Micetich, R. (1996) Modern Drug Design and Discovery: An Overview. *Curr Med Chem* **3:** 317–324.
Shah, K., Liu, Y., Deirmengian, C., and Shokat, K.M. (1997)

Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc Natl Acad Sci USA* **94:** 3565–3570.

Shen, K., Keng, Y.F., Wu, L., Guo, X.L., Lawrence, D.S., and Zhang, Z.Y. (2001) Acquisition of a Specific and Potent PTP1B Inhibitor from a Novel Combinatorial Library and Screening Procedure. *J Biol Chem* **276:** 47311–47319.

da Silva, M.L., Mortara, R.A., Barros, H.C., de Souza, W., and Trabulsi, L.R. (1989) Aggregation of membrane-associated actin filaments following localized adherence of enteropathogenic *Escherichia coli* to HeLa cells. *J Cell Sci* **93:** 439–446.

Sin, N., Meng, L., Wang, M.Q., Wen, J.J., Bornmann, W.G., and Crews, C.M. (1997) The anti-angiogenic agent fumagillin covalently binds and inhibits the methionine aminopeptidase, MetAP-2. *Proc Natl Acad Sci USA* **94:** 6099–6103.

Soete, M., and Dubremetz, J.F. (1996) *Toxoplasma gondii*: kinetics of stage-specific protein expression during tachyzoite-bradyzoite conversion *in vitro*. *Curr Topics Microb Immunol* **219:** 75–80.

Stockwell, B.R. (2000) Chemical genetics: ligand-based discovery of gene function. *Nat Rev Genet* **1:** 116–125.

Stockwell, B.R., Haggarty, S.J., and Schreiber, S.L. (1999) High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem Biol* **6:** 71–83.

Tallarico, J.A., Depew, K.M., Pelish, H.E., Westwood, N.J., Lindsley, C.W., Shair, M.D. *et al.* (2001) An alkylsilyl-tethered, high-capacity solid support amenable to diversity-oriented synthesis for one-bead, one-stock solution chemical genetics. *J Comb Chem* **3:** 312–318.

Tan, D.S., and Burbaum, J.J. (2000) Ligand discovery using encoded combinatorial libraries. *Current Opin Drug Discov Dev* **3(4):** 439–453.

Tanaka, H., Ohshima, N., and Hidaka, H. (1999) Isolation of cDNAs encoding cellular drug-binding proteins using a novel expression cloning procedure: drug-western. *Mol Pharmacol* **55:** 356–363.

Taraska, T., Ward, D.M., Ajioka, R.S., Wyrick, P.B., Davis-Kaplan, S.R., Davis, C.H., and Kaplan, J. (1996) The late chlamydial inclusion membrane is not derived from the endocytic pathway and is relatively deficient in host proteins. *Infect Immun* **64:** 3713–3727.

Tian, S.S., Lamb, P., King, A.G., Miller, S.G., Kessler, L., Luengo, J.I. *et al.* (1998) A small, nonpeptidyl mimic of granulocyte-colony-stimulating factor. *Science* **281:** 257–259.

Tilney, L.G., and Portnoy, D.A. (1989) Actin filaments and the growth, movement, and spread of the intracellular bacterial parasite, *Listeria monocytogenes*. *J Cell Biol* **109:** 1597–1608.

Walling, L.A., Peters, N.R., Horn, E.J., and King, R.W. (2002) New technologies for chemical genetics. *J Cell Biochem* **37:** 7–12.

Ward, G.E., and Carey, K.L. (1999) 96-well plates providing high optical resolution for high-throughput immunofluorescence-based screening of monoclonal antibodies against *Toxoplasma gondii*. *J Immunol Meth* **230:** 11–18.

Weinstein, J.N., Myers, T.G., O'Connor, P.M., Friend, S.H., Fornace, A.J. Jr, Kohn, K.W. *et al.* (1997) An information-intensive approach to the molecular pharmacology of cancer. *Science* **275:** 343–349.

Werbovetz, K.A. (2000) Target-based drug discovery for malaria, leishmaniasis, and trypanosomiasis. *Curr Med Chem* **7:** 835–860.

Yamaguchi, K., Yada, M., Tsuji, T., Hatanaka, Y., Goda, K., and Kobori, T. (1999) 4-Phenylthiazole derivatives inhibit IL-6 secretion in osteoblastic cells and suppress bone weight loss in ovariectomized mice. *Bioorg Med Chem Lett* **9:** 957–960.

Yarrow, J., Westwood, N., and Mitchison, T. (2000) A small molecule approach to cell motility. *Mol Biol Cell* **11:** 382a.

Zhang, B., Salituro, G., Szalkowski, D., Li, Z., Zhang, Y., Royo, I. *et al.* (1999) Discovery of a small molecule insulin mimetic with antidiabetic activity in mice. *Science* **284:** 974–977.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P. *et al.* (2001) Global analysis of protein activities using proteome chips. *Science* **293:** 2101–2105.